VOL. LXI. No. 242.] [April, 1952

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—THE IDENTITY OF INDISCERNIBLES

BY MAX BLACK

A. The principle of the Identity of Indiscernibles seems to me obviously true. And I don't see how we are going to define identity or establish the connexion between mathematics and logic without using it.

B. It seems to me obviously false. And your troubles as a mathematical logician are beside the point. If the principle is false you have no right to use it.

A. You simply *say* it's false—and even if you said so three times that wouldn't make it so.

B. Well, you haven't done anything more yourself than assert the principle to be true. As Bradley once said, "assertion can demand no more than counter-assertion; and what is affirmed on the one side, we on the other can simply deny".

A. How will this do for an argument? If two things, *a* and *b*, are given, the first has the property of being identical with *a*. Now *b* cannot have this property, for else *b* would be *a*, and we should have only one thing, not two as assumed. Hence *a* has at least one property, which *b* does not have, that is to say the property of being identical with *a*.

B. This is a roundabout way of saying nothing, for "*a* has the property of being identical with *a*" means no more than "*a* is *a*". When you begin to say "*a* is . . ." I am supposed to know what thing you are referring to as '*a*' and I expect to be told something about that thing. But when you end the sentence with the words ". . . is *a*" I am left still waiting. The sentence "*a* is *a*" is a useless tautology.

A. Are you as scornful about difference as about identity? For *a* also has, and *b* does not have, the property of being different from *b*. This is a second property that the one thing has but not the other.

B. All you are saying is that *b* is different from *a*. I think the form of words "*a* is different from *b*" does have the advantage over "*a* is *a*" that it might be used to give information. I might learn from hearing it used that '*a*' and '*b*' were applied to different things. But this is not what you want to say, since you are trying to use the names, not mention them. When I already know what '*a*' and '*b*' stand for, "*a* is different from *b*" tells me nothing. It, too, is a useless tautology.

A. I wouldn't have expected you to treat 'tautology' as a term of abuse. Tautology or not, the sentence has a philosophical use. It expresses the necessary truth that different things have at least one property not in common. Thus different things must be discernible; and hence, by contraposition, indiscernible things must be identical. Q.E.D.

B. Why obscure matters by this old-fashioned language? By "indiscernible" I suppose you mean the same as "having all properties in common". Do you claim to have proved that two things having all their properties in common are identical?

A. Exactly.

B. Then this is a poor way of stating your conclusion. If *a* and *b* are identical, there is just one thing having the two names '*a*' and '*b*'; and in that case it is absurd to say that *a* and *b* are two. Conversely, once you have supposed there are *two* things having all their properties in common, you can't without contradicting yourself say that *they* are "identical".

A. I can't believe you were really misled. I simply meant to say it is logically impossible for two things to have all their properties in common. I showed that *a* must have at least two properties—the property of being identical with *a*, and the property of being different from *b*—neither of which can be a property of *b*. Doesn't this prove the principle of Identity of Indiscernibles?

B. Perhaps you have proved something. If so, the nature of your proof should show us exactly what you have proved. If you want to call "being identical with *a*" a "property" I suppose I can't prevent you. But you must then accept the consequences of this way of talking. All you mean when you say "*a* has the property of being identical with *a*" is that *a* is *a*. And all you mean when you say "*b* does not have the property of being identical with *a*" is that *b* is not *a*. So what you have

"proved" is that *a* is *a* and *b* is not *a*, that is to say, *b* and *a* are different. Similarly, when you said that *a*, but not *b*, had the property of being different from *b*, you were simply saying that *a* and *b* were different. In fact you are merely redescribing the hypothesis that *a* and *b* are different by calling it a case of "difference of properties". Drop the misleading description and your famous principle reduces to the truism that different things are different. How true! And how uninteresting!

A. Well, the properties of identity and difference may be uninteresting, but they *are* properties. If I had shown that grass was green, I suppose you would say I hadn't shown that grass was coloured.

B. You certainly would not have shown that grass had any colour *other than* green.

A. What it comes to is that you object to the conclusion of my argument *following* from the premise that *a* and *b* are different.

B. No, I object to the triviality of the conclusion. If you want to have an interesting principle to defend, you must interpret "property" more narrowly—enough so, at any rate, for "identity" and "difference" not to count as properties.

A. Your notion of an interesting principle seems to be one which I shall have difficulty in establishing. Will you at least allow me to include among "properties" what are sometimes called "relational characteristics"—like *being married to Caesar* or *being at a distance from London?*

B. Why not? If you are going to defend the principle, it is for you to decide what version you wish to defend.

A. In that case, I don't need to count identity and difference as properties. Here is a different argument that seems to me quite conclusive. The only way we can discover that two different things exist is by finding out that one has a quality not possessed by the other or else that one has a relational characteristic that the other hasn't.

If *both* are blue and hard and sweet and so on, and have the same shape and dimensions and are in the same relations to everything in the universe, it is logically impossible to tell them apart. The supposition that in such a case there might really be two things would be unverifiable *in principle.* Hence it would be meaningless.

B. You are going too fast for me.

A. Think of it this way. If the principle were false, the fact that I can see only two of your hands would be no proof that you had just two. And even if every conceivable test agreed with

the supposition that you had two hands, you might all the time have three, four, or any number. You might have nine hands, different from one another and all indistinguishable from your left hand, and nine more all different from each other but indistinguishable from your right hand. And even if you really did have just two hands, and no more, neither you nor I nor anybody else could ever know that fact. This is too much for me to swallow. This is the kind of absurdity you get into, as soon as you abandon verifiability as a test of meaning.

B. Far be it from me to abandon your sacred cow. Before I give you a direct answer, let me try to describe a counter-example.

Isn't it logically possible that the universe should have contained nothing but two exactly similar spheres? We might suppose that each was made of chemically pure iron, had a diameter of one mile, that they had the same temperature, colour, and so on, and that nothing else existed. Then every quality and relational characteristic of the one would also be a property of the other. Now if what I am describing is logically possible, it is not impossible for two things to have all their properties in common. This seems to me to *refute* the Principle.

A. Your supposition, I repeat, isn't verifiable and therefore can't be regarded as meaningful. But supposing you *have* described a possible world, I still don't see that you have refuted the principle. Consider one of the spheres, *a*, . . .

B. How can I, since there is no way of telling them apart? *Which* one do you want me to consider?

A. This is very foolish. I mean either of the two spheres, leaving you to decide which one you wished to consider. If I were to say to you "Take any book off the shelf" it would be foolish on your part to reply "Which?"

B. It's a poor analogy. I know how to take a book off a shelf, but I don't know how to identify one of two spheres supposed to be alone in space and so symmetrically placed with respect to each other that neither has any quality or character the other does not also have.

A. All of which goes to show as I said before, the unverifiability of your supposition. Can't you imagine that one sphere has been designated as '*a*'?

B. I can imagine only what is logically possible. Now it is logically possible that somebody should enter the universe I have described, see one of the spheres on his left hand and proceed to call it '*a*'. I can imagine that all right, if that's enough to satisfy you.

A. Very well, now let me try to finish what I began to say about *a* . . .

B. I still can't let you, because you, in your present situation, have no right to talk about *a*. All I have conceded is that if something were to happen to introduce a change into my universe, so that an observer entered and could see the two spheres, one of them could then have a name. But this would be a different supposition from the one I wanted to consider. My spheres don't yet have names. If an observer were to enter the scene, he could perhaps put a red mark on one of the spheres. You might just as well say " By '*a*' I mean the sphere which would be the first to be marked by a red mark if anyone were to arrive and were to proceed to make a red mark ! " You might just as well ask me to consider the first daisy in my lawn that would be picked by a child, if a child were to come along and do the picking. This doesn't now distinguish any daisy from the others. You are just pretending to use a name.

A. And I think you are just pretending not to understand me. All I am asking you to do is to think of one of your spheres, no matter which, so that I may go on to say something about it when you give me a chance.

B. You talk as if naming an object and then thinking about it were the easiest thing in the world. But it isn't so easy. Suppose I tell you to name any spider in my garden : if you can catch one first or describe one uniquely you can name it easily enough. But you can't pick one out, let alone " name " it, by just thinking. You remind me of the mathematicians who thought that talking about an Axiom of Choice would really allow them to choose a single member of a collection when they had no criterion of choice.

A. At this rate you will never give me a chance to say anything. Let me try to make my point without using names. Each of the spheres will surely differ from the other in being at some distance from that other one, but at no distance from itself—that is to say, it will bear at least one relation to itself—*being at no distance from*, or *being in the same place as*—that it does not bear to the other. And this will serve to distinguish it from the other.

B. Not at all. *Each* will have the relational characteristic *being at a distance of two miles*, say, *from the centre of a sphere one mile in diameter*, etc. And each will have the relational characteristic (if you want to call it that) of *being in the same place as itself*. The two are alike in this respect as in all others.

A. But look here. Each sphere occupies a different place ; and this at least will distinguish them from one another.

B. This sounds as if you thought the places had some independent existence, though I don't suppose you really think so. To say the spheres are in "different places" is just to say that there is a distance between the two spheres; and we have already seen that will not serve to distinguish them. Each is at a distance—indeed the same distance—from the other.

A. When I said they were at different places I didn't mean simply that they were at a distance from one another. That one sphere is in a certain place does not entail the existence of any *other* sphere. So to say that one sphere is in its place, and the other in its place, and then to add that these places are different seems to me different from saying the spheres are at a distance from one another.

B. What does it mean to say "a sphere is in its place"? Nothing at all, so far as I can see. Where else could it be? *All* you are saying is that the spheres are in different places.

A. Then my retort is, What does it mean to say "Two spheres are in different places"? Or, as you so neatly put it, "Where else could they be?"

B. You have a point. What I should have said was that your assertion that the spheres occupied different places said nothing at all, unless you were drawing attention to the necessary truth that different physical objects must be in different places. Now if two spheres must be in different places, as indeed they must, to say that the spheres occupy different places is to say no more than they are two spheres.

A. This is like a point you made before. You won't allow me to deduce anything from the supposition that there are two spheres.

B. Let me put it another way. In the two-sphere universe, the only reason for saying that the places occupied were different would be that different things occupied them. So in order to show the places were different you would first have to show, in some other way, that the spheres were different. You will never be able to distinguish the spheres by means of the places they occupy.

A. A minute ago, you were willing to allow that somebody might give your spheres different names. Will you let me suppose that some traveller has visited your monotonous "universe" and has named one sphere "Castor" and the other "Pollux"?

B. All right—provided you don't try to use those names yourself.

A. Wouldn't the traveller, at least, have to recognise that *being at a distance of two miles from Castor* was not the same property as being at a distance of two miles *from Pollux?*

B. I don't see why. If he were to see that Castor and Pollux had exactly the same properties, he would see that "being at a distance of two miles from Castor" meant exactly the same as "being at a distance of two miles from Pollux".

A. They couldn't mean the same. If they did, "*being at a distance of two miles from Castor and at the same time not being at a distance of two miles from Pollux*" would be a self-contradictory description. But plenty of bodies could answer to this description. Again if the two expressions meant the same, anything which was two miles from Castor would have to be two miles from Pollux—which is clearly false. So the two expressions don't mean the same and the two spheres have at least two properties not in common.

B. Which?

A. *Being at a distance of two miles from Castor* and *being at a distance of two miles from Pollux.*

B. But now you are *using* the words "Castor" and "Pollux" as if they really stood for something. They are just our old friends '*a*' and '*b*' in disguise.

A. You surely don't want to say that the arrival of the name-giving traveller creates spatial properties? Perhaps we can't name your spheres and therefore can't name the corresponding properties; but the properties must be there.

B. What can this mean? The traveller has not visited the spheres, and the spheres have no names—neither 'Castor', nor 'Pollux', nor '*a*', nor '*b*', nor any others. Yet you still want to say they have certain properties which cannot be referred to without using names for the spheres. You want to say "the property of being at a distance from Castor" though it is logically impossible for you to talk in this way. You can't speak, but you won't be silent.

A. How eloquent, and how unconvincing! But since you seem to have convinced yourself, at least, perhaps you can explain another thing that bothers me: I don't see that you have a right to talk as you do about places or spatial relations in connexion with your so-called "universe". So long as we are talking about our own universe—*the* universe—I know what you mean by "distance", "diameter", "place" and so on. But in what you want to call a universe, even though it contains only two objects, I don't see what such words could mean. So far as I can see, you are applying these spatial terms in their present usage to a hypothetical situation which contradicts the presuppositions of that usage.

B. What do you mean by "presupposition"?

A. Well, you spoke of measured distances, for one thing. Now this presupposes some means of measurement. Hence your "universe" must contain at least a third thing—a ruler or some other measuring device.

B. Are you claiming that a universe must have at least three things in it? What is the least number of things required to make a world?

A. No, all I am saying is that you cannot describe a configuration as *spatial* unless it includes at least three objects. This is part of the meaning of "spatial"—and it is no more mysterious than saying you can't have a game of chess without there existing at least thirty-five things (thirty-two pieces, a chessboard, and two players).

B. If this is all that bothers you, I can easily provide for three or any number of things without changing the force of my counter-example. The important thing, for my purpose, was that the configuration of two spheres was symmetrical. So long as we preserve this feature of the imaginary universe, we can now allow any number of objects to be found in it.

A. You mean any *even* number of objects.

B. Quite right. Why not imagine a plane running clear through space, with everything that happens on one side of it always exactly duplicated at an equal distance in the other side.

A. A kind of cosmic mirror producing real images.

B. Yes, except that there wouldn't be any mirror! The point is that in *this* world we can imagine any degree of complexity and change to occur. No reason to exclude rulers, compasses, and weighing machines. No reason, for that matter, why the Battle of Waterloo shouldn't happen.

A. Twice over, you mean—with Napoleon surrendering later in two different places simultaneously!

B. Provided you wanted to call both of them "Napoleon".

A. So your point is that everything could be duplicated on the other side of the non-existent Looking Glass. I suppose whenever a man got married, his identical twin would be marrying the identical twin of the first man's fiancée?

B. Exactly.

A. Except that "identical twins" wouldn't be *numerically* identical?

B. You seem to be agreeing with me.

A. Far from it. This is just a piece of gratuitous metaphysics. If the inhabitants of your world had enough sense to know what was sense and what wasn't, they would never suppose all the events in their world were duplicated. It would be much more

sensible for them to regard the "second" Napoleon as a mere mirror image—and similarly for all the other supposed "duplicates".'

B. But they could walk through the "mirror" and find water just as wet, sugar just as sweet, and grass just as green on the other side.

A. You don't understand me. They would not postulate "another side". A man looking at the "mirror" would be seeing *himself*, not a duplicate. If he walked in a straight line toward the "mirror" he would eventually find himself back at his starting point, not at a duplicate of his starting point. This would involve their having a different geometry from ours—but that would be preferable to the logician's nightmare of the reduplicated universe.

B. They might think so—until the twins really began to behave differently for the first time!

A. Now it's you who are tinkering with your supposition. You can't have your universe and change it too.

B. All right, I retract.

A. The more I think about your "universe" the queerer it seems. What would happen when a man crossed your invisible "mirror"? While he was actually crossing, his body would have to change shape, in order to preserve the symmetry. Would it gradually shrink to nothing and then expand again?

B. I confess I hadn't thought of that.

A. And here is something that explodes the whole notion. Would you say that one of the two Napoleons in your universe had his heart in the right place—literally, I mean?

B. Why, of course.

A. In that case his "mirror-image" twin would have the heart on the opposite side of the body. One Napoleon would have his heart on the left of his body, and the other would have it on the right of his body.

B. It's a good point, though it would still make objects like spheres indistinguishable. But let me try again. Let me abandon the original idea of a *plane* of symmetry and to suppose instead that we have only a *centre* of symmetry. I mean that everything that happened at any place would be exactly duplicated at a place an equal distance on the opposite side of the centre of symmetry. In short, the universe would be what the mathematicians call "radially symmetrical". And to avoid complications we could suppose that the centre of symmetry itself was physically inaccessible, so that it would be impossible for any material body to pass through it. Now in *this* universe,

identical twins would have to be either both right-handed or both left-handed.

A. Your universes are beginning to be as plentiful as blackberries. You are too ingenious to see the force of my argument about verifiability. Can't you see that your supposed description of a universe in which everything has its "identical twin" doesn't describe anything verifiably different from a corresponding universe without such duplication? This must be so, no matter what kind of symmetry your universe manifested.

B. You are assuming that in order to verify that there are two things of a certain kind, it must be possible to show that one has a property not possessed by the other. But this is not so. A pair of very close but similar magnetic poles produce a characteristic field of force which assures me that there are two poles, even if I have no way of examining them separately. The presence of two exactly similar stars at a great distance might be detected by some resultant gravitational effect or by optical interference—or in some such similar way—even though we had no way of inspecting one in isolation from the other. Don't physicists say something like this about the electrons inside an atom? We can verify *that* there are two, that is to say a certain property of the whole configuration, even though there is no way of detecting any character that uniquely characterises any element of the configuration.

A. But if you were to approach your two stars one would have to be on your left and one on the right. And this would distinguish them.

B. I agree. Why shouldn't we say that the two stars are distinguishable—meaning that it would be possible for an observer to see one on his left and the other on his right, or more generally, that it would be *possible* for one star to come to have a relation to a third object that the second star would not have to that third object.

A. So you agree with me after all.

B. Not if you mean that the two stars do not have all their properties in common. All I said was that it was logically possible for them to enter into different relationships with a third object. But this would be a change in the universe.

A. If you are right, nothing unobserved would be observable. For the presence of an observer would always change it, and the observation would always be an observation of something else.

B. I don't say that every observation changes what is observed. My point is that there isn't any *being to the right* or *being to the*

left in the two-sphere universe until an observer is introduced, that is to say until a real change is made.

A. But the spheres themselves wouldn't have changed.

B. Indeed they would: they would have acquired new relational characteristics. In the absence of any asymmetric observer, I repeat, the spheres would have all their properties in common (including, if you like, the power to enter into different relations with other objects). Hence the principle of Identity of Indiscernibles is false.

A. So perhaps you really do have twenty hands after all?

B. Not a bit of it. Nothing that I have said prevents me from holding that we can verify *that* there are exactly two. But we could know *that* two things existed without there being any way to distinguish one from the other. The Principle is false.

A. I am not surprised that you ended in this way, since you assumed it in the description of your fantastic "universe". Of course, if you began by assuming that the spheres were numerically different though qualitatively alike, you could end by "proving" what you first assumed.

B. But I wasn't "proving" anything. I tried to support my contention that it is logically possible for two things to have all their properties in common by giving an illustrative description. (Similarly, if I had to show it is logically possible for nothing at all to be seen I would ask you to imagine a universe in which everybody was blind.) It was for you to show that my description concealed some hidden contradiction. And you haven't done so.

A. All the same I am not convinced.

B. Well, then, you ought to be.[1]

Cornell University.

[1] The following notes and references might be helpful to anybody wishing to make up his own mind on the questions raised:

The Definition of Identity: See *Principia Mathematica*, vol. i, definition 13.01. The theory of types required Whitehead and Russell to say x and y are identical if and only if the same predicative functions are satisfied by both. For a similar definition see W. V. Quine, *Mathematical Logic* (1940), definition D 10 (p. 136). See also G. Frege, *The Foundations of Arithmetic* (Oxford, 1940, p. 76).

Self-evidence of the Principle: "I think it is obvious that the principle of Identity of Indiscernibles is not true" (G. E. Moore, *Philosophical Studies*, 1922, p. 307). "Leibniz's 'principles of indiscernibles' is all nonsense. No doubt, all things differ; but there is no logical necessity for it" (C. S. Peirce, *Collected Papers*, 4.311). "Russell's definition of '=' won't do; because according to it one cannot say that two objects have all their properties in common (even if this proposition is never true,

it is nevertheless *significant*)." (L. Wittgenstein, *Tractatus Logico-Philosophicus*, 5.5302). See also C. H. Langford, "Otherness and dissimilarity" in MIND, vol. xxxix (1930), pp. 454-461.

Bradley's remark: Not made in connexion with *this* topic. See *Ethical Studies*, 2nd edn., 1927, p. 165. Bradley affirmed what he called "the Axiom of the Identity of Indiscernibles" (*The Principles of Logic*, 2nd ed., 1928, p. 288).

Identity as a relational property: ". . . numerical identity, which is a dyadic relation of a subject to itself of which nothing but an existent individual is capable" (C. S. Peirce, *Collected Papers*, 1.461).

The proof of the principle by treating identity as a property: "It should be observed that by 'indiscernibles' he [Leibniz] cannot have meant two objects which agree as to *all* their properties, for one of the properties of x is to be identical with x, and therefore this property would necessarily belong to y if x and y agreed in *all* their properties. Some limitation of the common properties necessary to make things indiscernible is therefore implied by the necessity of an axiom" (*Principia Mathematica*, vol. i, p. 51). *Cf.* K. Grelling, "Identitas indiscernibilium" in *Erkenntnis*, vol. vi (1936), pp. 252-259.

Counter-examples: C. D. Broad tries to refute McTaggart's form of the principle ("The dissimilarity of the diverse") by the example of a universe consisting of two minds, without bodies, that are exactly alike in all respects. (*Examination of McTaggart's Philosophy*, vol. i, p. 176). Broad holds however, that "either spatial or temporal separation involves dissimilarity" (*op. cit.*, p. 173).

The argument from verifiability: "To say that B and C are 'really' two, although they seem one, is to say something which, if B and C are totally indistinguishable, seems wholly devoid of meaning" (B. Russell, *An Inquiry into Meaning and Truth*, 1940, p. 127).

The distinguishability of asymmetric bodies and their mirror images: There is a famous discussion of this in Kant's *Critique of Pure Reason*. See, for instance, H. Vaihinger, *Kommentar zur Kant's Kritik der reinen Vernunft*, 1922, vol. ii, pp. 518-532 ("Anhang – Das Paradoxon der symmetrischen Gegenstände").

II.—INTERPRETATIONS OF MODAL LOGIC

By G. H. von Wright

I

Some of the notorious difficulties in connexion with modal concepts concern the relation of what we shall call higher order modalities to lower order modalities.

It is a "received truth" of modal logic that if a proposition is true, then it is also possibly true. (*Ab esse ad posse valet consequentia.*) By an application of this principle to the case when the proposition concerned is to the effect that a further proposition is possibly true or not possibly true, we conclude:

If (it is true that) a proposition is possibly true, then the proposition in question is also possibly possibly true, and if (it is true that) a proposition is not possibly true, then the proposition in question is also possibly not possibly true.

It cannot be a universally valid truth of modal logic that if a proposition is possibly true, then it is (certainly) true. In the special case, however, when the possibly true proposition is to the effect that a further proposition is possibly true, it might be thought that possibility entails truth. In other words: it might be suggested that, if a proposition is possibly possibly true, then (it is true that) the proposition in question is possibly true.

In view of what was concluded above from the *ab esse ad posse* principle, this suggestion would entail an identity between what is possible and what is possibly possible. We shall refer to this as the First Identity.

It cannot be a universally valid truth of modal logic that if a proposition is possibly false, then it is (certainly) false. In the special case, however, when the possibly false proposition is to the effect that a further proposition is possibly true, it might be thought that possible falsehood entails falsehood. In other words, it might be suggested that, if a proposition is possibly not possibly true, then (it is true that) the proposition in question is not possibly true.

In view of what was concluded above from the *ab esse ad posse* principle, this suggestion would entail an identity between what is impossible (not possible) and what is possibly impossible. We shall refer to this as the Second Identity.

Possibility and necessity are interdefinable. What is necessarily true is what is not possibly false, and *vice versa*. The reader can easily satisfy himself that, if we prefer to speak in terms of necessity rather than in terms of possibility, then our First Identity becomes an identity between what is necessary and what is necessarily necessary, and our Second Identity becomes an identity between what is not necessary and what is necessarily not necessary.

It can be shown that the First Identity is weaker than the Second. This means: We can accept the First Identity without accepting the Second, but we cannot accept the Second Identity without also accepting the First. (*Cf.* below, Sections V and VII).

The two identities are consistent with what we shall call the "received system" of modal logic. The notion of a "received system", however, is not clear in itself. We might identify it with Lewis's *S2*. I shall, however, here identify it with a more comprehensive system, which includes *S2* and which differs substantially from *S2* only in that it allows us to conclude from any given theorem of the system to the necessary truth of that theorem. I call this system "received", because I do not think that anyone would seriously question the necessary and universal validity of its theorems for the notions of possibility and necessity.[1]

The identity of what is possible with what is possibly possible is the constitutive feature of the system of modal logic known as *S4*, and the identity of what is impossible with what is possibly impossible is the constitutive feature of *S5*.

The test of truth for the two identities is thus not a test of consistency with other, accepted truths. The question then arises:

On what grounds, if on any, can it be decided whether the two identities are valid, or not?

Our "logical intuitions", apparently, give no strong indication in favour of any definite answer to the question of truth here. One of the main reasons for this, it seems to me, is the fact that higher order modal expressions like "possibly possible" or "possibly impossible" have hardly any *use* at all in ordinary or scientific discourse (outside modal logic).

A problem of primary importance, therefore, is to *invent a use* or some kind of "equivalent" of a use for the expressions in question (outside modal logic). This done, the problem of

[1] A more detailed study of this system will be found in my publication *An Essay in Modal Logic*, Amsterdam, 1951.

truth mentioned above can be tackled on a firmer basis. We shall return to this problem in the concluding section of the present paper.

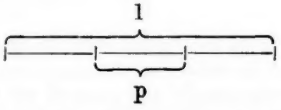
II

We shall first consider a geometrical interpretation of ("classical") propositional logic.

The variables of the calculus are p, q, r, . . . (an unlimited multitude). The constants which we are going to use are $\sim$ for negation, & for conjunction, v for disjunction, $\rightarrow$ for material implication, and $\longleftrightarrow$ for material equivalence. The class of well-formed formulae (wff : s) is defined recursively in the usual way. Rules for brackets are given.

We consider an arbitrary line of unit length 1. All wff : s can be represented by segments of the unit line in accordance with the following rules of interpretation :

i. The variables are represented by arbitrary segments. (A segment may consist of parts which are disconnected.) For example :



ii. The negation of a wff. is represented by that part of the unit line which is not covered by the segment representing the wff. itself. We shall call the segment representing the negation of a wff. the *complement* of the segment representing the wff. itself.

iii. The conjunction of two wff : s is represented by the common part, if any, of the segments representing the wff : s themselves.

It follows from these rules that the disjunction of two wff : s is represented by the part of the unit line covered jointly by the segments representing the wff : s themselves ; that the implication of an antecedent and a consequent wff. is represented by the part of the unit line covered jointly by the segments representing the consequent wff. and the negation of the antecedent wff. ; and that the equivalence of two wff : s is represented by the part of the unit line covered jointly by the segments representing the conjunction of the two wff : s and the conjunction of the negations of the two wff : s.

A wff. is a tautology of propositional logic, if and only if its representation in accordance with Rules *i-iii* is the unit line,

independently of the choice of segments to represent the variables. For example: The segment representing $p \vee \sim p$ is the part covered jointly by an arbitrary segment of the unit line (representing p) and its complement. Hence the representation of $p \vee \sim p$ must be the unit line itself.

If and only if the representation of the implication of two wff:s is the unit line, the representation of the antecedent is a segment of the representation of the consequent. For suppose, *e.g.*, that the representation of $p \rightarrow q$ is the unit line. This means that the representation of $\sim p$ covers at least those portions of the unit line which are not covered by the representation of q. Hence the representation of p "falls inside", *i.e.* is a segment of, the representation of q.

If and only if the representation of the equivalence of two wff:s is the unit line, the wff:s are represented by the same segment.

As a name of the unit line itself we shall use the symbol 1. As a name of the segment representing the negation of a wff., the representation of which is the unit line, we shall use the symbol 0.

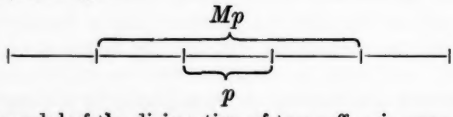
III

To the calculus of propositional logic we add one new constant M. It is prefixed to wff:s in the same was as $\sim$ is prefixed. We modify the definition of the class of wff:s so as to include mention of M. No new rules for brackets are needed.

Just as we by the negation of a given wff. understand the wff. which we get, when $\sim$ is prefixed to the given wff., we shall by the *modal* of a given wff. understand the wff. which we get, when M is prefixed to the given wff.

All wff:s can be represented as segments of a unit line, if —in addition to the three rules of the previous section—we adopt the following new rules of interpretation:

iv. The modal of a wff. is represented by an arbitrary segment of the unit line which covers the segment representing the wff. itself. For example:



v. The modal of the disjunction of two wff:s is represented by that part of the unit line which is covered jointly by the segments representing the modals of the two wff:s themselves.

vi. If two wff:s are represented by the same segment, then the

modals of the two wff : s will have to be represented by the same segment as well.

vii. The representation of the modal of a wff., the representation of which is 0, is 0.

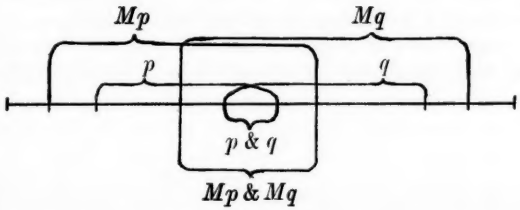
A wff. will be called a *M*-tautology or modal tautology, if and only if its representation in accordance with Rules *i-vii* is the unit line, independently of the choice of segments to represent the variables and the modals.

(Since the wff : s now under consideration include among themselves the wff : s of propositional logic, it follows that the modal tautologies include among themselves the tautologies of propositional logic).

Let us convince ourselves, relying upon the geometrical interpretation of wff : s, that $M(p \,\&\, q) \rightarrow Mp \,\&\, Mq$ is a *M*-tautology. ("If propositions are mutually consistent, they are self-consistent.")

Remembering what was said above about the representation of implications, the task can be reduced to one of showing that the representation of $M\ (p \,\&\, q)$ must be a segment of the representation of $Mp \,\&\, Mq$.

We start from the picture



and ask ourselves the following question : How is the representation of $M(p \,\&\, q)$ to be fitted into this picture ?

The segment representing p is the same as the segment representing $p \,\&\, q \vee p \,\&\, \sim q$. According to Rule *vi*, the modal of p must then be represented by the same segment as the modal of $p \,\&\, q \vee p \,\&\, \sim q$. But in virtue of Rule *v*, the modal of $p \,\&\, q \vee p \,\&\, \sim q$ is represented by the same segment as the representation of the disjunction of the modal of $p \,\&\, q$ and the modal of $p \,\&\, \sim q$. It follows that the representation of $M(p \,\&\, q)$ must be a segment of the representation of Mp, which is in the picture.

By an exactly similar argument, it is shown that the representation of $M\ (p \,\&\, q)$ must also be a segment of the representation of Mq, which is in the picture.

The two conditions, thus deduced, which the representation of $M(p \,\&\, q)$ has to satisfy, can be simultaneously fulfilled if and only if the representation of $M(p \,\&\, q)$ is a segment of the representation of $Mp \,\&\, Mq$, which is in the picture.

A second example will be given. Let us convince ourselves that $M(p \rightarrow \sim M \sim p)$ is a M-tautology.

The segment representing $p \rightarrow \sim M \sim p$ is the same as the segment representing $\sim p \vee \sim M \sim p$. According to Rule *vi*, the modal of $p \rightarrow \sim M \sim p$ must then be represented by the same segment as the modal of $\sim p \vee \sim M \sim p$. But in virtue of Rule *v*, the modal of $\sim p \vee \sim M \sim p$ is represented by the same segment as the representation of the disjunction of the modal of $\sim p$ and the modal of $\sim M \sim p$.

The representation of $M \sim p \vee \sim M \sim p$ is the unit line. Since, according to Rule *iv*, the representation of $\sim M \sim p$ is a segment of the representation of $M \sim M \sim p$, it follows that the representation of $M \sim p \vee M \sim M \sim p$ must be the unit line too. *Q.E.D.*

For any given wff. of the class under consideration can be effectively decided, whether or not it is a M-tautology. The proof will not be given here.

The class of M-tautologies includes within itself the class of theorems of *S2*. The proof will not be given here.

The class of M-tautologies is identical with the class of what is called in this paper the " received truths " of modal logic.

IV

We consider the same class of wff : s as in the previous section, but add a new rule for the interpretation of modals :

viii. The representation of the modal of the modal of a wff. is the same as the representation of the modal of the wff.

The rule answers to what we called above the First Identity (what is possible = what is possibly possible).

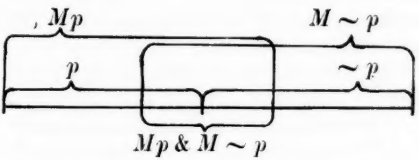
A wff. will be called a M'-tautology, if and only if its representation in accordance with Rules *i-viii* is the unit line, independently of the choice of segments to represent the variables and the modals.

(It follows trivially that M-tautologies are also M'-tautologies.)

Let us convince ourselves, relying upon the geometrical interpretation of wff : s, that $M(Mp \,\&\, M \sim p) \longleftrightarrow Mp \,\&\, M \sim p$ is a M'-tautology. (" The contingent is the same as the possibly contingent.")

Remembering what was said above about the representation of equivalences, the task can be reduced to one of showing that the representation of $M(Mp \,\&\, M \sim p)$ must be the same segment as the representation of $Mp \,\&\, M \sim p$.

We start from the picture



and ask ourselves the following question : How is the representation of $M(Mp \,\&\, M \sim p)$ to be fitted into this picture ?

We know from what was shown in the preceding section that the representation of $M(Mp \,\&\, M \sim p)$ must be a segment of the representation of $MMp \,\&\, MM \sim p$. In virtue of Rule *viii*, however, the representation of MMp is the same segment as the representation of Mp, and the representation of $MM \sim p$, the same segment as the representation of $M \sim p$. Consequently, the representation of $MMp \;\&\; MM \sim p$ is the same segment as the representation of $Mp \,\&\, M \sim p$, which is already in the picture.

We have now shown that the representation of $M(Mp \,\&\, M \sim p)$ must be a segment of the representation of $Mp \,\&\, M \sim p$. In virtue of Rule *iv*, on the other hand, the representation of $Mp \,\&\, M \sim p$ must be a segment of the representation of $M(Mp \,\&\, M \sim p)$. These two features of the representation of $M(Mp \,\&\, M \sim p)$ can be reconciled only if the representation of $M(Mp \,\&\, M \sim p)$ in the picture is the same segment as the representation of $Mp \,\&\, M \sim p$.

For any given wff. of the class under consideration can be effectively decided, whether or not it is a M'-tautology. The proof will not be given here.

The class of M'-tautologies is identical with the class of theorems in *S4*.

V

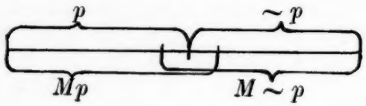
We consider the same class of wff : s as in Sections III and IV, but add a further rule :

ix. The representation of the modal of the negation of the modal of a wff. is the same as the representation of the negation of a modal of the wff.

This rule answers to what we called above the Second Identity (what is impossible = what is possibly impossible).

A wff. will be called a *M''*-tautology, if and only if its representation in accordance with Rules *i-vii* and *ix* is the unit line, independently of the choice of segments to represent the variables and the modals.

It can be shown that *M'*-tautologies are also *M''*-tautologies. Consider the picture



We ask the question: How is the representation of the modal of the modal of *p*, *i.e.* *MMp*, to be fitted into the picture?

In virtue of Rule *ix*, the representation of the modal of the negation of the modal of *p*, *i.e.* $M \sim Mp$, is the same as the representation of the negation of the modal of *p*, *i.e.*, $\sim Mp$, which is in the picture. In virtue of Rule *vi*, if two wff:s have the same representation, then their modals have the same representation too. Thus the representation of $M \sim M \sim Mp$ if the same as the representation of the modal of the modal of *p*, *i.e.* *MMp*. But in virtue of Rule *ix*, the representation of $M \sim M \sim Mp$ is the same as the representation of $\sim M \sim Mp$ which was already shown to be the same as the representation of *Mp*. Thus we conclude that the representation of the modal of the modal of *p*, *i.e.* *MMp*, is the same segment as the representation of the modal of *p*, *i.e.* *Mp*, which is already in the picture.

For any given wff. of the class under consideration can be effectively decided, whether or not it is a *M''*-tautology. The proof will not be given here.

The class of *M''*-tautologies is identical with the class of theorems in *S5*.

VI

We shall now consider physical models which are "isomorphic" with the above geometrical models of propositional logic and various systems of modal logic.

We conceive of the variables *p*, *q*, *r*, . . . as the "physical" parts of some "physical" whole. (The parts need not necessarily be spatio-temporally continuous.) For example: *p*, *q*, *r*, . . . can be represented by segments of a metal rod, geographical areas on the earth's surface, populations of human beings, etc.

All the wff : s of propositional logic can be interpreted as parts of such wholes, if we adopt the rule that the negation of a wff. is represented by that part of the whole which remains, when the part representing the wff. itself has been removed, *and* the rule that the conjunction of two wff : s is represented by the common part, if any, of the parts representing the wff : s themselves.

Consider next some " activity " or " process " which may occur in the parts of a whole under consideration. For example : an electric current of a certain strength running through a segment of a metal rod, rainfall in a district, a certain disease occurring among (not necessarily all) members of a population, etc.

Relative to such an assumed activity, we decide to call the parts *fields of activity.*

The activity under consideration may have some " influence " or " effect ", extending from a part in which the activity is supposed to occur to other parts of the same whole. For example : an electric current running through a segment of a metal rod may have a thermic effect on the surrounding portions of the rod ; a change of climate in one country may produce a change of climate in neighbouring countries ; a disease among members of one population may spread by contagion to the members of another population, etc.

Relative to an assumed activity and an assumed influence, we decide to call the part, to which the influence may extend, when the activity occurs in a given part, the *field of influence* of (or associated with) the given field of activity.

It will be convenient to understand the concept of a field of influence in such a way that any given field of influence *includes* the associated field of activity.

Relative to an assumed activity and an assumed influence, we further decide to call the part, to which the influence will *not* extend, when the activity occurs in a given part, the *field of resistance* of the surrounding (of the given field of activity).

The activities and influences to be considered are assumed to obey the following three laws :

L1. If a field of activity is divided into parts (not necessarily mutually exclusive), then the fields of influence of the parts conceived of as fields of the same activity, are jointly identical with the field of influence of the original field of activity.

L2. With the same field of activity is always associated the same field of influence.

L3. If the field of activity is non-existent (" the 0-field "), then the associated field of influence is also non-existent.

Activities which obey *L1* can conveniently be called *additive* with regard to their influences. Far from all processes in nature possess this feature of additiveness. But some obviously possess it. For example: If an epidemic occurs among the members of two populations, then the members of the surrounding population which may become infected are exactly those which may get the disease from members of the one *or* the other of the two populations, whenever the disease occurs in them.

We can now give a new interpretation to the class of wff: s considered in Sections III-V (the wff: s formed of the symbols of propositional logic and the symbol *M*).

We call the parts of a given whole fields of activity. The variables, negation, and conjunction are interpreted as before. The modal of a wff. is interpreted as the field of influence of the field of activity representing the wff. itself.

(Since, under the terminology adopted, any part of the given whole is called a field of activity, it follows that the fields of influence (and of resistance) are a sub-class of the fields of activity.)

A wff. is a *M*-tautology, if and only if its representation is the whole, of which the fields of activity are parts, independently of the choice of fields of activity to represent the variables and fields of influence to represent the modals.

VII

Processes in nature which are additive with regard to effects of theirs in the sense of *L1* above, can be divided into the following three categories:

A. Processes which are such that, if a field of influence of theirs becomes itself a field of activity for the same process, then the field of influence of this new field of activity extends beyond the boundaries of the original field of influence.

B. Processes which are such that, if a field of influence of theirs becomes itself a field of activity for the same process, then the field of influence of this new field of activity does *not* extend beyond the boundaries of the original field of influence.

C. Processes which are such that, if a field of resistance of the neighbourhood of a field of activity of theirs becomes itself a field of activity for the same process, then the field of influence of this new field of activity does not extend beyond the boundaries of the original field of resistance.

Processes of the third kind are also processes of the second kind, but not necessarily *vice versa*.

The fields of activity, influence, and resistance of processes of the kind *B* offer a model of a system of modal logic in which the First Identity holds.

The fields of activity, influence, and resistance of processes of the kind *C* offer a model of a system of modal logic in which the Second Identity holds.

Examples of processes of each of the three kinds can be given. Here we shall consider only one example.

Let us conceive of p as a certain population of human beings. Let the activity, of which p is a field, be an epidemic disease. (We need not suppose, however, that all members of the population actually have or even can get the disease.) $\sim p$ is represented by the population, of which a man is a member if and only if he is not a member of the population representing p. For the sake of convenience, we shall speak of the "population" and its "neighbourhood" (population).

Of Mp we conceive as the population itself *and* those members of the neighbourhood, who for some reason or other may catch the disease from members of the population. The representation of Mp can thus conveniently be called the field of influence of the epidemic in the population.

It follows that $\sim Mp$ is represented by those members of the neighbourhood, who for some reason or other cannot catch the disease from members of the population. The representation of $\sim Mp$ can thus conveniently be called the field of resistance ("immunity") of the neighbourhood.

If to the field of influence are added those members, if any, of the neighbourhood who, although they cannot catch the epidemic from members of the population, yet may for some reason or other catch it from such of their fellow members of the neighbourhood as may themselves catch it from members of the population, we get a representation of MMp.

If to the field of resistance are added those members, if any, of the field of influence who, for some reason or other, may catch the disease from a member of the field of resistance, we get a representation of $M \sim Mp$.

As to the relative magnitude of the population representing Mp and MMp, two different situations might occur:

1. The members of the field of resistance are totally immune from the disease and cannot, consequently, catch it from anybody. In this case, the populations representing Mp and MMp coincide.

2. The members of the field of resistance are secure from catching the disease from members of the population and yet not

secure from catching it from infected fellow members of theirs of the neighbouring population. (This could be the case for biological reasons, but also because, say, the members of the field of resistance are segregated from contact with members of the population.) In this case, the population representing MMp would be larger than the population representing Mp.

If the fields of resistance of all populations concerned consist entirely of members, if any, which cannot catch the disease in question from non-members of those fields, we obtain a model of a modal logic, in which the First Identity holds.

As to the relative magnitude of the populations representing $\sim Mp$ and $M \sim Mp$, two different situations might also occur:

1. The members of the field of influence are secure from catching the disease from members of the field of resistance. (This would trivially be the case, if the members of the field of resistance are totally immune and, consequently, unable to acquire the disease. It would non-trivially be the case, *e.g.*, if members of the field of resistance could acquire the disease, though only in a non-catching form.) Under this alternative the populations representing $\sim Mp$ and $M \sim Mp$ coincide.

2. The members of the field of influence are not secure from catching the disease from members of the field of resistance. (It may, for example, happen that a member of the field of resistance catches the disease from a fellow member of the neighbourhood population, who has himself caught it from a member of the population, and that subsequently another member of the field of influence catches the disease from this infected member of the field of resistance.) Under this alternative, the population representing $M \sim Mp$ would be larger than the population representing $\sim Mp$.

If the fields of influence of all populations concerned consist entirely of members, if any, who cannot catch the disease in question from non-members of those fields, we obtain a model of a modal logic, in which the Second Identity holds.

It is easy to see that a model in which the Second Identity holds (universally) is also a model in which the First Identity holds (universally). If the members of the population representing $\sim Mp$ cannot infect the members of the population representing Mp, then the second population equals its own field of resistance and the first its own field of influence. But if it is universally true of fields of influence that they consist entirely of members who cannot catch the disease from non-members, it follows that the population representing MMp must coincide with the population representing Mp.

VIII

We have seen that the "logical behaviour" of natural processes which are, in a defined sense, *additive* with regard to their effects is governed by the same formal rules as the concept of possibility. These processes—or their fields of activity, influence, and resistance—can be said to afford models or illustrations of the possible. In the language which speaks about these processes there is a use for formal equivalents of higher order modalities and thus also for formal equivalents of the two identities with which we have been concerned in this paper. The criteria of truth at hand are such as to show that for some additive natural processes neither of the identities holds, for some both identities hold, and for some only the first identity holds.

It seems to me that the existence of such illustrations as these among natural processes shows that neither of the identities in question can claim universal validity for the concept of the possible (and the necessary) in modal logic, nor can either of them be rejected as contrary to the "true" nature of modality.

It would be of some interest to investigate the "logical behaviour" of non-additive natural processes with a view to the question whether this behaviour also illustrates some general notions of logic.

Cambridge University.

III.—KNOWLEDGE AND BELIEF

By Norman Malcolm

"We must recognize that when we know something we either do, or by reflecting, can know that our condition is one of knowing that thing, while when we believe something, we either do or can know that our condition is one of believing and not of knowing: so that we cannot mistake belief for knowledge or vice versa."[1]

This remark is worthy of investigation. Can I discover *in myself* whether I know something or merely believe it?

Let us begin by studying the ordinary usage of "know" and "believe". Suppose, for example, that several of us intend to go for a walk and that you propose that we walk in Cascadilla gorge. I protest that I should like to walk beside a flowing stream and that at this season the gorge is probably dry. Consider the following cases:

(1) You say "I believe that it won't be dry although I have no particular reason for thinking so". If we went to the gorge and found a flowing stream we should not say that you *knew* that there would be water but that you thought so and were right.

(2) You say "I believe that it won't be dry because it rained only three days ago and usually water flows in the gorge for at least that long after a rain". If we found water we should be inclined to say that you knew that there would be water. It would be quite natural for you to say "I knew that it wouldn't be dry"; and we should tolerate your remark. This case differs from the previous one in that here you had a *reason*.

(3) You say "I know that it won't be dry" and give the same reason as in (2). If we found water we should have very little hesitation in saying that you knew. Not only had you a reason, but you *said* "I know" instead of "I believe". It may seem to us now that the latter should not make a difference—but it does.

(4) You say "I know that it won't be dry" and give a stronger reason, *e.g.* "I saw a lot of water flowing in the gorge when I passed it this morning". If we went and found water, there would be no hesitation at all in saying that you knew. If, for example, we later met someone who said "Weren't you surprised to see water in the gorge this afternoon?" you would reply "No,

[1] H. A. Prichard, *Knowledge and Perception* (Oxford, 1950), p. 88.

I *knew* that there would be water; I had been there earlier in the day". We should have no objection to this statement.

(5) Everything happens as in (4), except that upon going to the gorge we find it to be dry. We should not say that you knew, but that you *believed* that there would be water. And this is true even though you declared that you knew, and even though your evidence was the same as it was in case (4) in which you did know.

I wish to make some comments on the usage of "know", "knew", "believe", and "believed", as illustrated in the preceding cases:

(*a*) Whether we should say that you knew, depends in part on whether you had grounds for your assertion and on the strength of those grounds. There would certainly be less hesitation to say that you knew in case (4) than in case (3), and this can be due only to the difference in the strength of the grounds.

(*b*) Whether we should say that you knew, depends in part on how *confident* you were. In case (2), if you had said "It rained only three days ago and usually water flows in the gorge for at least that long after a rain; but, of course, I don't feel absolutely sure that there will be water", then we should *not* have said that you knew that there would be water. If you lack confidence that p is true then others do not say that you know that p is true, even though *they* know that p is true. Being confident is a necessary condition for knowing.

(*c*) Prichard says that if we reflect we cannot mistake belief for knowledge. In case (4) you knew that there would be water, and in case (5) you merely believed it. Was there any way that you could have discovered by reflexion, in case (5), that you did not know? It would have been useless to have reconsidered your grounds for saying that there would be water, because in case (4), where you *did* know, your grounds were identical. They could be at fault in (5) only if they were at fault in (4), and they were not at fault in (4). Cases (4) and (5) differ in only one respect—namely, that in one case you did subsequently find water and in the other you did not. Prichard says that we can determine by reflexion whether we know something or merely believe it. But where, in these cases, is the material that reflexion would strike upon? There is none.

There is only one way that Prichard could defend his position. He would have to say that in case (4) you did *not* know that there would be water. And it is obvious that he would have said this. But this is false. It is an enormously common usage of language to say, in commenting upon just such an incident as (4), "He knew that the gorge wouldn't be dry because he had seen water

flowing there that morning". It is a usage that all of us are familiar with. We so employ "know" and "knew" every day of our lives. We do not think of our usage as being loose or incorrect—and it is not. As philosophers we may be surprised to observe that it *can* be that the knowledge that *p* is true should differ from the belief that *p* is true *only* in the respect that in one case *p* is true and in the other false. But that is the fact.

There is an argument that one is inclined to use as a proof that you did not know that there would be water. The argument is the following: It could have turned out that you found no water; if it had so turned out you would have been mistaken in saying that you would find water; therefore you could have been mistaken; but if you could have been mistaken then you did not know.

Now it certainly *could* have turned out that the gorge was quite dry when you went there, even though you saw lots of water flowing through it only a few hours before. This does not show, however, that you did not know that there would be water. What it shows is that *although you knew you could have been mistaken.* This would seem to be a contradictory result; but it is not. It seems so because our minds are fixed upon another usage of "know" and "knew"; one in which "It could have turned out that I was mistaken", implies "I did not know".

When is "know" used in this sense? I believe that Prichard uses it in this sense when he says that when we go through the proof of the proposition that the angles of a triangle are equal to two right angles we *know* that the proposition is true (p. 89). He says that if we put to ourselves the question: Is our condition one of knowing this, or is it only one of being convinced of it?, then "We can only answer 'Whatever may be our state on other occasions, here we are knowing this.' And this statement is an expression of our *knowing* that we are knowing; for we do not *believe* that we are knowing this, we know that we are" (p. 89). He goes on to say that if someone were to object that we may be making a mistake "because for all we know we can later on discover some fact which is incompatible with a triangle's having angles that are equal to two right angles, we can answer that we *know* that there can be no such fact, for in knowing that a triangle must have such angles we also know that nothing can exist which is incompatible with this fact" (p. 90).

It is easy to imagine a non-philosophical context in which it would have been natural for Prichard to have said "I know that the angles of a triangle are equal to two right angles". Suppose that a young man just beginning the study of geometry was in

doubt as to whether that proposition is true, and had even constructed an ingenious argument that appeared to prove it false. Suppose that Prichard was unable to find any error in the argument. He might have said to the young man: "There must be an error in it. I know that the angles of a triangle are equal to two right angles."

When Prichard says that "nothing can exist which is incompatible with" the truth of that proposition, is he prophesying that no one will ever have the ingenuity to construct a flawless-looking argument against it? I believe not. When Prichard says that "we" *know* (and implies that *he* knows) that the proposition is true and *know* that nothing can exist that is incompatible with its being true, he is not making any *prediction* as to what the future will bring in the way of arguments or measurements. On the contrary, he is asserting that *nothing* that the future might bring could ever count as evidence against the proposition. He is implying that he would not *call* anything "evidence" against it. He is using "know" in what I shall call its "strong" sense. "Know" is used in this sense when a person's statement "I know that p is true" implies that the person who makes the statement would look upon nothing whatever as evidence that p is false.

It must not be assumed that whenever "know" is used in connexion with mathematical propositions it is used in the strong sense. A great many people have *heard* of various theorems of geometry, *e.g.* the Pythagorean. These theorems are a part of "common knowledge". If a schoolboy doing his geometry assignment felt a doubt about the Pythagorean theorem, and said to an adult, "Are you *sure* that it is true?", the latter might reply "Yes, I know that it is". He might make this reply even though he could not give proof of it and even though he had never gone through a proof of it. If subsequently he was presented with a "demonstration" that the theorem is false, or if various persons reputed to have a knowledge of geometry soberly assured him that it is false, he might be filled with doubt or even be convinced that he was mistaken. When he said "Yes, I know that it is true", he did not pledge himself to hold to the theorem through thick and thin. He did not absolutely exclude the possibility that something could prove it to be false. I shall say that he used "know" in the "weak" sense.

Consider another example from mathematics of the difference between the strong and weak senses of "know". I have just now rapidly calculated that 92 times 16 is 1472. If I had done this in the commerce of daily life where a practical problem was

at stake, and if someone had asked "Are you sure that $92 \times 16 = 1472$?", I might have answered "I *know* that it is; I have just now calculated it". But also I might have answered "I know that it is; but I will calculate it again to *make sure*". And here my language points to a distinction. I say that I *know* that $92 \times 16 = 1472$. Yet I am willing to *confirm* it—that is, there is something that I should *call* "making sure"; and, likewise, there is something that I should *call* "finding out that it is false". If I were to do this calculation again and obtain the result that $92 \times 16 = 1372$, and if I were to carefully check this latter calculation without finding any error, I should be disposed to say that I was previously mistaken when I declared that $92 \times 16 = 1472$. Thus when I say that I know that $92 \times 16 = 1472$, I allow for the possibility of a *refutation*; and so I am using "know" in its weak sense.

Now consider propositions like $2 + 2 = 4$ and $7 + 5 = 12$. It is hard to think of circumstances in which it would be natural for me to say that I know that $2 + 2 = 4$, because no one ever questions it. Let us try to suppose, however, that someone whose intelligence I respect argues that certain developments in arithmetic have shown that $2 + 2$ does not equal 4. He writes out a proof of this in which I can find no flaw. Suppose that his demeanour showed me that he was in earnest. Suppose that several persons of normal intelligence became persuaded that his proof was correct and that $2 + 2$ does not equal 4. What would be my reaction? I should say "I can't see what is wrong with your proof; but it *is* wrong, because I *know* that $2 + 2 = 4$. Here I should be using "know" in its strong sense. I should not admit that any argument or any future development in mathematics could show that it is false that $2 + 2 = 4$.

The propositions $2 + 2 = 4$ and $92 \times 16 = 1472$ do not have the same status. There *can* be a demonstration that $2 + 2 = 4$. But a demonstration would be for me (and for any average person) only a curious exercise, a sort of *game*. We have no serious interest in proving that proposition.[1] It does not *need* a proof. It stands without one, and would not fall if a proof went against it. The case is different with the proposition that $92 \times 16 = 1472$. We take an interest in the demonstration (calculation) because

[1] Some logicians and philosophers have taken an interest in proving that $2 + 2 = 4$ (e.g. Leibnitz, *New Essays on the Understanding*, Bk. IV, Ch. vii, Sec. 10; Frege, *The Foundations of Arithmetic*, §6). They have wished to show that it can be deduced from certain premises, and to determine what premises and rules of inference are required in the deduction. Their interest has not been in the *outcome* of the deduction.

that proposition *depends* upon its demonstration. A calculation may lead me to reject it as false. But $2 + 2 = 4$ does *not* depend on its demonstration. It does not depend on anything! And in the calculation that proves that $92 \times 16 = 1472$, there are steps that do not depend on any calculation (*e.g.* $2 \times 6 = 12$; $5 + 2 = 7$; $5 + 9 = 14$).

There is a correspondence between this dualism in the logical status of mathematical propositions and the two senses of "know". When I use "know" in the weak sense I am prepared to let an investigation (demonstration, calculation) determine whether the something that I claim to know is true or false. When I use "know" in the strong sense I am not prepared to look upon anything as an *investigation;* I do not concede that anything whatsoever could prove me mistaken; I do not regard the matter as open to any *question;* I do not admit that my proposition could turn out to be false, that any future investigation *could* refute it or cast doubt on it.[1]

We have been considering the strong sense of "know" in its application to mathematical propositions. Does it have application anywhere in the realm of *empirical* propositions—for example, to propositions that assert or imply that certain physical things exist? Descartes said that we have a "moral assurance" of the truth of some of the latter propositions but that we lack a "metaphysical certainty".[2] Locke said that the perception of the existence of physical things is not "so certain as our intuitive knowledge, or the deductions of our reason" although "it is an assurance that deserves the name of knowledge".[3] Some philosophers have held that when we make judgments of perception such as that there are peonies in the garden, cows in the field, or dishes in the cupboard, we are "taking for granted" that the peonies, cows, and dishes exist, but not knowing it in the "strict" sense. Others have held that all empirical propositions, including judgments of perception, are merely hypotheses.[4] The

[1] Compare these remarks about the strong sense of "know" with some of Locke's statements about "intuitive knowledge": ". . . in this the mind is at no pains of proving or examining. . . ." "This part of knowledge . . . leaves no room for hesitation, doubt, or examination. . . ." "It is on this intuition that depends all the certainty and evidence of all our knowledge; which certainly every one finds to be so great, that he cannot imagine, and therefore not require a greater. . . ." Locke, *Essay*, Bk. IV, Ch. ii, §1.

[2] Descartes, *Discourse on the Method*, Part IV.

[3] Locke, *Essay*, Book IV, Ch. xi, §3.

[4] *E.g.* ". . . no proposition, other than a tautology, can possibly be anything more than a probable hypothesis." Ayer, A. J., *Language, Truth and Logic*, p. 38 (2nd edn.).

thought behind this exaggerated mode of expression is that any empirical proposition whatever *could* be refuted by future experience—that is, it *could* turn out to be false. Are these philosophers right?

Consider the following propositions:

(i) The sun is about ninety million miles from the earth.
(ii) There is a heart in my body.
(iii) Here is an ink-bottle.

In various circumstances I should be willing to assert of each of these propositions that I know it to be true. Yet they differ strikingly. This I see when, with each, I try to imagine the possibility that it is false.

(i) If in ordinary conversation someone said to me "The sun is about 20 million miles from the earth, isn't it?" I should reply "No; it is about 90 million miles from us". If he said "I think that you are confusing the sun with Polaris", I should reply, "I *know* that 90 million miles is roughly the sun's distance from the earth". I might invite him to verify the figure in an encyclopedia. A third person who overheard our conversation could quite correctly report that I knew the distance to the sun, whereas the other man did not. But this knowledge of mine is little better than hearsay. I have seen that figure mentioned in a few books. I know nothing about the observations and calculations that led astronomers to accept it. If tomorrow a group of eminent astronomers announced that a great error had been made and that the correct figure is 20 million miles, I should not insist that they were wrong. It would surprise me that such an enormous mistake could have been made. But I should no longer be willing to say that I *know* that 90 million is the correct figure. Although I should *now* claim that I know the distance to be about 90 million miles, it is easy for me to envisage the possibility that some future investigation will prove this to be entirely false.

(ii) Suppose that after a routine medical examination the excited doctor reports to me that the X-ray photographs show that I have no heart. I should tell him to get a new machine. I should be inclined to say that the fact that I have a heart is one of the few things that I can count on as absolutely certain. I can feel it beat. I know it's there. Furthermore, how could my blood circulate if I didn't have one? Suppose that later on I suffer a chest injury and undergo a surgical operation. Afterwards the astonished surgeons solemnly declare that they searched my chest cavity and found no heart, and that they made

incisions and looked about in other likely places but found it not. They are convinced that I am without a heart. They are unable to understand how circulation can occur or what accounts for the thumping in my chest. But they are in agreement and obviously sincere, and they have clear photographs of my interior spaces. What would be my attitude? Would it be to insist that they were all mistaken? I think not. I believe that I should eventually accept their testimony and the evidence of the photographs. I should consider to be false what I now regard as an absolute certainty.

(ii) Suppose that as I write this paper someone in the next room was to call out to me "I can't find an ink-bottle; is there one in the house"? I should reply "Here is an ink-bottle". If he said in doubtful tone "Are you sure? I looked there before", I should reply "Yes, I know there is; come and get it".

Now could it turn out to be false that there is an ink-bottle directly in front of me on this desk? Many philosophers have thought so. They would say that many things could happen of such a nature that if they did happen it would be proved that I am deceived. I agree that many extraordinary things could happen, in the sense that there is no logical absurdity in the supposition. It could happen that when I next reach for this ink-bottle my hand should seem to pass *through* it and I should not feel the contact of any object. It could happen that in the next moment the ink-bottle will suddenly vanish from sight; or that I should find myself under a tree in the garden with no ink-bottle about; or that one or more persons should enter this room and declare with apparent sincerity that they see no ink-bottle on this desk; or that a photograph taken now of the top of the desk should clearly show all of the objects on it except the ink-bottle. Having admitted that these things *could* happen, am I compelled to admit that if they did happen then it would be proved that there is no ink-bottle here *now?* Not at all! I could say that when my hand seemed to pass through the ink-bottle I should *then* be suffering from hallucination; that if the ink-bottle suddenly vanished it would have miraculously ceased to exist; that the other persons were conspiring to drive me mad, or were themselves victims of remarkable concurrent hallucinations; that the camera possessed some strange flaw or that there was trickery in developing the negative. I admit that in the next moment I could find myself under a tree or in the bathtub. But this is not to admit that it could be revealed in the next moment that I am now dreaming. For what I admit is that I might be instantaneously transported to the garden, but not that

in the next moment I might *wake up* in the garden. There is nothing that could happen to me in the next moment that I should call "waking up"; and therefore nothing that could happen to me in the next moment would be accepted by me now as proof that I now dream.

Not only do I not *have* to admit that those extraordinary occurrences would be evidence that there is no ink-bottle here; the fact is that I *do not* admit it. There is nothing whatever that could happen in the next moment or the next year that would by me be called *evidence* that there is not an ink-bottle here now. No future experience or investigation could prove to me that I am mistaken. Therefore, if I were to say "I know that there is an ink-bottle here", I should be using "know" in the strong sense.

It will appear to some that I have adopted an *unreasonable* attitude toward that statement. There is, however, nothing unreasonable about it. It seems so because one thinks that the statement that here is an ink-bottle *must* have the same status as the statements that the sun is 90 million miles away and that I have a heart and that there will be water in the gorge this afternoon. But this is a *prejudice*.

In saying that I should regard nothing as evidence that there is no ink-bottle here now, I am not *predicting* what I should do if various astonishing things happened. If other members of my family entered this room and, while looking at the top of this desk, declared with apparent sincerity that they see no ink-bottle, I might fall into a swoon or become mad. I *might* even come to believe that there is not and has not been an ink-bottle here. I cannot foretell with certainty how I should react. But if it is *not* a prediction, what is the meaning of my assertion that I should regard nothing as evidence that there is no ink-bottle here?

That assertion describes my *present* attitude towards the statement that here is an ink-bottle. It does not prophesy what my attitude *would* be if various things happened. My present attitude toward that statement is radically different from my present attitude toward those other statements (*e.g.* that I have a heart). I do *now* admit that certain future occurrences would disprove the latter. Whereas no imaginable future occurrence would be considered by me *now* as proving that there is not an ink-bottle here.

These remarks are not meant to be autobiographical. They are meant to throw light on the common concepts of evidence, proof, and disproof. Everyone of us upon innumerable occasions of

daily life takes this same attitude towards various statements about physical things, *e.g.* that here is a torn page, that this dish is broken, that the thermometer reads 70, that no rug is on the floor. Furthermore, the concepts of proof, disproof, doubt, and conjecture, *require* us to take this attitude. In order for it to be possible that any statements about physical things should *turn out to be false* it is necessary that some statements about physical things *cannot* turn out to be false.

This will be made clear if we ask ourselves the question, When do we *say* that something turned out to be false ? ; When do we use those words ? Someone asks you for a dollar. You say " There is one in this drawer ". You open the drawer and look, but it is perfectly empty. Your statement turned out to be false. This can be said because you *discovered* an empty drawer. It could not be said if it were only probable that the drawer is empty or were still open to question. Would it make sense to say " I had better make sure that it is empty ; perhaps there is a dollar in it after all ? " Sometimes ; but not always. Not if the drawer lies open before your eyes. That remark is the prelude to a search. What search can there be when the emptiness of the drawer confronts you ? In certain circumstances there is nothing that you would call " making sure " that the drawer is empty ; and likewise nothing that you would call " its turning out to be false " that the drawer is empty. You *made* sure that the drawer is empty. One statement about physical things *turned out to be false* only because you *made sure* of another statement about physical things. The two concepts cannot exist apart. Therefore it is impossible that *every* statement about physical things *could* turn out to be false.

In a certain important respect some *a priori* statements and some empirical statements possess the same logical character. The statements that $5 \times 5 = 25$ and that here is an ink-bottle, both lie beyond the reach of doubt. On both my judgment and reasoning *rests*. If you could somehow undermine my confidence in either, you would not teach me *caution*. You would fill my mind with chaos ! I could not even make *conjectures* if you took away those fixed points of certainty ; just as a man cannot *try* to climb whose body has no support. A conjecture implies an understanding of what certainty would be. If it is not a certainty that $5 \times 5 = 25$ and that here is an ink-bottle, then I do not understand what it is. You cannot make me doubt either of these statements or treat them as hypotheses. You cannot persuade me that future experience could refute them. With both of them it is perfectly unintelligible to me to speak of a

"possibility" that they are false. This is to say that I know both of them to be true, in the strong sense of "know". And I am inclined to think that the strong sense of "know" is what various philosophers have had in mind when they have spoken of "perfect", "metaphysical", or "strict certainty".[1]

It will be thought that I have confused a statement about my "sensations", or my "sense-data", or about the way something *looks* or *appears* to me, with a statement about physical things. It will be thought that the things that I have said about the statement "Here is an ink-bottle" could be true only if that statement is interpreted to mean something like "There appears to me to be an ink-bottle here", *i.e.* interpreted so as not to assert or imply that any physical thing exists. I wish to make it clear that my statement "Here is an ink-bottle" is *not* to be interpreted in that way. It would be utterly fantastic for me in my present circumstances to say "There appears to me to be an ink-bottle here". The usage of the latter sentence would be natural only in two sorts of circumstances: either when I know or believe that there is *not* an ink-bottle before me or when I am in doubt as to whether there is an ink-bottle before me. But I do not know or believe that there is not an ink-bottle here; nor am I in *doubt* as to whether there is—indeed, I *cannot* be in doubt, *i.e.* a doubt is unintelligible to me.

If someone were to call me on the telephone and say that he urgently needed an ink-bottle I should invite him to come here and get this one. If he said that it was extremely urgent that he should obtain one immediately and that he could not afford to waste time going to a place where there might not be one, I should tell him that it is an absolute certainty that there is one here, that nothing could be more certain, that it is something I absolutely guarantee. But if my statement "There is an ink-bottle here" were a statement about my "sensations" or "sense-data", or if it meant that there *appears* to me to be an ink-bottle here or that something here *looks* to me like an ink-bottle, and if that is all that I meant by it—then I should react quite differently to his urgent request. I should say that there is probably

[1] Descartes, for example, apparently took as his criterion for something's being "entirely certain" that he could not *imagine* in it the least ground of doubt: ". . . je pensai qu'il fallait . . . que je retasse comme absolument faux tout ce en quoi je pourrais imaginer le moindre doute, afin de voir s'il ne me resterait point après cela quelque chose en ma creánce qui fut entièrement indubitable". *Discourse*, Part IV. And Locke (as previously noted) said of "intuitive knowledge" that one *cannot imagine* a greater certainty, and that it "leaves no room for hesitation, doubt, or examination". *Essay*, Bk. IV, Ch. II, §1.

XUM

an ink-bottle here but that I could not *guarantee* it, and that if he needs one very desperately and at once then he had better look elsewhere. In short, I wish to make it clear that my statement " Here is an inkbottle " is strictly about physical things and not about " sensations ", " sense-data ", or " appearances ".

Going back to Prichard's remark that we can determine by reflexion whether we know something or merely believe it, we can conclude that with respect to one usage of " know " (weak) what he says is false, but with respect to another usage (strong) what he says is true. It is distinctive of the latter usage that when I say that I know that something is true, it has for me no meaning to speak of the possibility of its being proved false. Reflexion can teach me that I know something in this sense. It can make me aware of the fact that I should not *call* anything proof or evidence of its falsity. Prichard would say that we know (in the strong sense) various *a priori* truths, and in this I agreed. I further argued that in this same sense of " know " we know various empirical truths about physical things; which is something that Prichard did not say and would have thought wrong. I tried to show that the distinction between strong and weak knowledge does not run parallel with the distinction between *a priori* and empirical statements but cuts across it, *i.e.* we may distinguish these two senses of " know " *within a priori* statements and *within* empirical statements.

I wish to end by saying that Ludwig Wittgenstein, in discussion, gave me the principal idea of this paper—namely, that there is a resemblance in logic between some *a priori* and some empirical statements.

Cornell University.

IV.—PRAGMATISM, REALISM AND POSITIVISM IN THE UNITED STATES

By James Ward Smith

Three introductory comments will perhaps clarify my intention in this paper:

1. My principal point is systematic rather than historical. I have attempted to examine with some impartiality one segment in the line of twentieth-century philosophising, and to approach it from the perspective of the questions: What is going on here? What are people disagreeing about? Do the same labels always mean the same things, and do different labels always mean different things? Whenever I allow myself the luxury of this kind of questioning I find myself confronted by certain peculiar puzzles. This paper is an attempt to throw one such puzzle into focus.

2. I have selected certain historical material as my base, and I have chosen in particular certain movements which have had considerable vogue in the United States during recent decades. For those who may be unfamiliar with the ideas of men like Peirce, James and Dewey I have tried to state as accurately as I can their opinions on the issues relevant to this paper. I have not, however, made any attempt to offer pedagogical summaries of their philosophies.

3. I trust that there will be no objection to the word "Positivism". I have used it in an extremely broad sense, and use such phrases as "analytic movement" as synonymous with it.[1] The words "positivist" and "positivism" seem to be in general disfavour, largely, I suspect, as the result of incriminating overtones which accrued to them in the twenties and thirties. I prefer them, however, to such cumbersome alternatives as

[1] The unqualified phrase "analytic movement" will not serve to express my meaning, for in one sense of the word "analytic" certain so-called empiricists from Hume to H. H. Price may be called analytic philosophers, and this is not the sense of analysis which I here have in mind. What I do have in mind can only be described by historical reference to such groups as the Vienna Circle and the influence exerted upon analytic philosophy by certain men who at one time accepted the labels "positivist" and "positivism". Though certain men whom I shall mention (like Quine and Popper) are seldom, if ever, called "positivists", the kind of analysis in which they indulge is the direct result of tendencies in philosophy fostered by such groups as the Vienna Circle. It is these tendencies I wish to discuss.

XUM

"logical empiricism", not only because of their simplicity, but also because if cleansed of abusive overtones they are in a general sort of way quite appropriate.

Pragmatist Theories of Meaning and Truth

I begin by distinguishing between the three very different senses in which Peirce, James and Dewey used the word "Pragmatism". This distinction depends primarily upon the difference between the task of specifying the conditions under which a word or a concept or an idea can be said to have meaning, and the quite different, if closely related, task of specifying the conditions under which a proposition can be said to be true. To offer a rough approximation at the outset: The pragmatic maxim is proposed by Peirce as a rule for the attainment of "clarity of our conceptions", *i.e.* as a theory of meaning, and is associated with an account of the notion of truth which is not itself appropriately labelled "pragmatic". James, brilliant philosopher though he in many respects was, constantly muddled theory of meaning and theory of truth. He accordingly stated the principle of pragmatism alternately as a theory of the nature of meaning and a theory of the nature of truth. Dewey, more rigorous than James, if less clear, preferred to eliminate the concept of truth. In its stead he substituted a concept of "warranted assertibility" which, in conjunction with Peirce's pragmatic account of meaning, seemed to Dewey an adequate explication of the conditions of successful thinking.

Stated less roughly, the essential differences between these three leading pragmatists are as follows:

Peirce was himself a physical scientist, and his philosophy is in large degree the result of speculation upon the methods of physical science. There seemed to him to be three major steps involved in scientific method upon which philosophers would do well to ponder. The first is the formulation of hypotheses as the result of observation. A scientist confronted by data does not just stare at them; he formulates an idea or an hypothesis. The second step is the eliciting from the hypothesis of the conditions of further experiment. A scientist in possession of an hypothesis does not just stand in awe of it. He determines the conditions of future observations (*i.e.* of his experiments) by deducing what further things would be the case if the hypothesis were assumed to be correct. The third step consists in the actual experimenting and the inevitable alteration of the hypothesis in the light of new data. In short, the three-fold process begins again and continues indefinitely.

Peirce, speculating upon this method, came to a great many rather startling conclusions. But among them there are two which are central, and which form the base upon which everything else he said rests. I am interested solely in these two central principles. Unless one understands both, one will constantly misinterpret Peirce.

The first principle is the Principle of Pragmatism. It relates almost exclusively to what I have described as the second step of scientific method. Suppose that a scientist is asked to clarify the meaning of any hypothesis—or, what comes to the same thing, of the concepts contained in it. Clearly the only sensible thing he can do is precisely to " elicit from it the conditions of further experiment ". When the scientist deduces from his hypothesis the sensible consequences which would be observed if the hypothesis were correct (*i.e.* when he sets up his experiment) precisely what he is doing is to clarify the meaning of his hypothesis. This and this only is the import of Peirce's pragmatic maxim. " Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then our conception of these effects is the whole of our conception of the object."

Peirce was obsessed with the idea that traditional notions of clarification by definition put, as he liked to say, the cart in front of the horse. You cannot give a clear definition of a concept which is fuzzy. You must clarify your meaning before you can hope for a clear definition. Peirce formulated his pragmatic maxim in an attempt to get the horse in front of the cart. And in so doing what he did was in effect to present the grandfather of all twentieth-century theories of the meaning of meaning.

I turn now, however, to the second fundamental principle. This Peirce called the " Principle of Fallibilism ", and it relates to a very different feature of the scientific method from that to which the pragmatic maxim relates. It springs from what Peirce held to be the intrinsically directional character of the method. It is an essential presupposition of the method we have described that the succession of constantly altered hypotheses is not only a succession but a progression. That is to say, the whole process will be indulged in (so Peirce insisted) only if it is assumed that it is directional. And to Peirce the assumption that a series is directional is equivalent to the assumption that it has an ideal limit. His label for the ideal limit toward which science must be assumed to be directed is the words " The Truth ".

The principle of fallibilism states unequivocally that one is never justified in assuming that the latest and best and so-far-

uncontroverted hypothesis *is* (or accurately expresses) this ideal limit. This is to say, that one must always distinguish the latest and best hypothesis—which is, one must assume, that hypothesis whose meaning is most fully clarified in accordance with the pragmatic maxim—one must, I say, always distinguish this from the Truth (often called the Real) which constitutes the ideal limit of the process. This means indeed that even if (in God's eye) you had hold of the Truth by the tail you could never be sure (that is to say, you could never know) that you did.

The conjunction of these two principles in Peirce underscores as nothing else can the fundamental character of his distinction between the problem of meaning and the problem of truth. The truth relation is repeatedly defined by Peirce in straightforward old-fashioned correspondence terms. He was a thorough-going realist, both about things and about concepts, and considered the question of truth as a question of the relation between propositions in our discourse to events occurring independently of that discourse. To define the conditions of *stable belief* (*i.e.* of clarity of conception and of the conditions to be fulfilled by the "latest and best hypothesis") is one thing: to define *truth* is another. The single definition of truth upon which later pragmatists have placed so much emphasis, that truth consists in future agreement of qualified observers, would never have caused the confusion it has caused had it been soberly read within the context of Peirce's larger and more basic doctrine of fallibilism. For consider what that more basic doctrine insists that all qualified observers must agree upon (whether they are "future" is irrelevant): they must agree that no opinion, though it be the latest and the best and the one unanimously asserted, can be assumed to be the last word—*i.e.* incorrigible—*i.e.* identical with the truth.

William James's pragmatism rests upon what is perhaps an understandable, but what is also and notwithstanding an indefensible *non sequitur*. To begin with, James constantly talks of the truth of concepts and of percepts rather than of propositions. And the key to his modification of Peirce is his unwavering conviction that to specify the conditions for a concept's having meaning is *eo ipso* to specify the conditions for that concept's being true. In short, if a concept literally means what you do with it, its truth must lie in a successful doing. If Peirce's distinctions are accepted, this Jamesian hypothetical is a dreadful muddle.

It is not my purpose, however, to prove that James was muddled. My sole purpose at the moment is to point out that,

rightly or wrongly, James made the equation which Peirce, in formulating the principle of fallibilism, refused to allow. James defines the truth in terms of the conditions to be satisfied by the latest and best hypothesis. A stable belief is true. In James it is impossible to distinguish, where in Peirce it is necessary to distinguish, between the two very different tasks : (*a*) of determining the criteria whereby we convince ourselves that an hypothesis is the reasonable one to adopt (*i.e.* achieve stable belief), and (*b*) of ascertaining the meaning of truth.

John Dewey, especially under attack by Russell, does something rather different. Dewey does not confuse conditions for stability of belief with the meaning (correspondence) of truth. What Dewey does rather is to avow that if by " true " we persist in meaning (as, of course, Russell does) something like " corresponds with the facts " (if in calling " Caesar crossed the Rubicon " " true " we mean that there is a relation of correspondence between a statement uttered now and an event which took place two thousand years ago) then we can never in this world (and there is for Dewey no next) determine whether any proposition is true or not. (Of course, Peirce said we could never be sure.) Dewey accordingly pleads that we discard this mysterious and unprofitable notion altogether, and substitute for it a conception of " warranted assertibility " which is in all major respects identical with Peirce's conception of the conditions to be fulfilled by the " latest and best hypothesis ". That is to say, where James equated two things which Peirce demanded we keep distinct, Dewey frankly and ruthlessly discards one of the two. Where in James pragmatism entails confusion between meaning and truth, in Dewey it entails rejection of the traditional conception of truth and its replacement by a corollary of the theory of meaning, which corollary is supposed to perform the legitimate function of the old notion without, as it were, committing its metaphysical nonsense.

Realism as Implied by Pragmatism

I wish now to turn to a second fundamental distinction between Peirce, James and Dewey. " Pragmatism " is not the only word which they use constantly but with a difference. " Realism " is another such word. Each of these men has called himself a " realist ". Yet " realist ", as I shall try to indicate, meant something quite different to each. To begin once more by stating my point very roughly : Peirce considered realism as an inescapable epistemological premiss, and not at all as a thesis

in metaphysics. He was a metaphysician, but it is not his metaphysics that he called "realism"; he considered his metaphysics a form of idealism. James, on the other hand, certainly conceived of realism as a metaphysic presupposed by his particular brand of pragmatism. Dewey, who unlike either Peirce or James rejects metaphysics, is in a separate camp altogether, and his realism functions as an implicit rather than an explicit feature of his system.

When Peirce lists alternative metaphysics, realism does not occur in the list. Having rejected all forms of pluralism, including dualism, he cites as the three possible monistic metaphysics: materialism, idealism and neutral monism. Of these he rejects the first and the last, and proceeds to offer his peculiar brand of pan-psychism as to him the most acceptable form of idealism. Realism, which is not itself a special metaphysic alternative to others, is simply for Peirce a basic epistemological premiss of all intelligent discourse.

There are two distinct facets to Peirce's realism. Each relates to a special feature in his analysis of scientific method. On the one hand, he considers it an inescapable presumption of that method that in experiment and observation we are dealing with that which is independent of us. This is realism, if you will, concerning particular occurrences and events. On the other hand, the directional character of the process, that feature which forces upon Peirce the principle of fallibilism, requires the presumption that the end result we are seeking is also independent of us and of our search for it. It is when interpreting this latter presumption that Peirce feels driven to the Platonic account of concepts in connexion with which he is so often discussed. If our method rests upon a presumed tendency for fact to conform to law, and if the results of applying the method confirm this presupposition, it followed for Peirce that law is precisely something in its own right and that the hypotheses in which we attempt to formulate laws do in fact "stand for" something "real". The principle of fallibilism warns us never to suppose that the "standing for" is a faithful reproduction. But laws are something in their own right, and objective universals are the "stuff" of which objective laws are made.

Since Peirce conceived of metaphysics as a "knowledge about", and since for him realism is no more than an inescapable "presumption that", it would have seemed to him nonsense to speak of realism as a metaphysic. In his sense, idealism, materialism and neutral monism are all equally "realistic". Materialism seemed to him patent nonsense, neutral monism a verbal evasion,

and idealism the only defensible attempt to hazard a theory about the real. If, however, his idealist metaphysic was an attempt to formulate an hypothesis about the nature of the real, his realism is itself no more than a methodological principle that there is a real to formulate hypotheses about. He freely admitted that realism is an unargued premiss, but insisted (*a*) that there is no alternative to it, and (*b*) that no amount of application of the method which employs it could possibly lead to doubt of it.

James's realism is of a wholly different sort. It is customary to assert that James's thought falls into three separate periods which are in the end unintegrated and incompatible. This is a mistake. It is true that James's thought falls into three discernible periods. In the beginning he is a psychologist, in the middle he is a pragmatist, in the end he is a realist. But for James all three stages are as it were peas in the same pod. And though I recognise the incomplete and fragmentary character of James's philosophical writings, I submit that we should recognise that in this he was right. His pragmatism is not only consistent with but clearly springs in detail from his functionalism in psychology. And his final realism is no more than the making explicit of certain fundamental principles implicit, though largely unexpressed, both in his psychology and in his pragmatism.

The point here, however, is that what James made explicit in his later realism was a positive metaphysic. It is indeed precisely that metaphysic of neutral monism which Peirce accused of verbal evasion.

James realised that to define what the mind is by describing what the mind does (*i.e.* the central point of his psychology) implies that the meaning of a concept is determined by the way in which it functions in discourse (*i.e.* the central point of his pragmatism). But he also realised that the specific descriptions of mental functioning which he offered in the psychology were in terms of neural arcs, laryngeal spasms, habits and pushes of effort, all of which were there assumed to be what they are independently of our observation of them. And he further realised that the relation of successful functioning which he repeatedly analysed in his pragmatism was assumed to take place in contexts which were objective, independent, and "real". What impressed him in particular was that, since the functional definition of mind in the psychology involved the rejection of any distinction in kind between the mental and the material, the objective independent contexts, with which he knew himself to be concerned all along, could in all consistency not be labelled (by him at any rate) as either mental or material. He therefore

faced the issue squarely, and in 1904 proclaimed his doctrine of neutral monism.

To trace the subsequent history of this doctrine in the United States is to study the movement of neo-Realism. I shall have a few remarks to make about this movement shortly. But before I continue with this aspect of the tale, I should like to call attention to the third way in which pragmatism has been associated with realism, namely, in the writings of Dewey.

Dewey, as I have said, rejects metaphysics altogether. This places him at once in a position rather closer to Peirce than to James. For in Dewey, as in Peirce, realism is not expounded as a metaphysical theory ; it is simply accepted as an epistemological principle. At this point, however, all similarity between Dewey and Peirce seems to me to end.

One must face the fact that Dewey's writing is exceedingly unclear. It is never possible to be entirely certain that one has hold of what Dewey intends to assert. It is quite certain, however, that in defending his conception of "warranted assertibility" against Russell's attack he rejects altogether the correspondence view of truth. Since for Peirce the realist presumption of science was, in effect, its commitment to the correspondence view of truth, it follows that the realist presumption in Dewey cannot be what it is in Peirce.

What I think Dewey has in mind is something like the following: There are problems, and there are investigators into problems, and the whole business of intellection is an actual form of behaviour indulged in by investigators involved with problems. In short, what Peirce described as the method of science is a form of behaviour actually occurrent in the world. Realism for Dewey means the facing of this fact, and the abandonment of all metaphysical presuppositions to the effect that there is anything more going on here than actual behaviour. Warranted opinion is one kind of behaviour, unwarranted opinion another kind of behaviour and the distinction between the two involves no appeal to an abstruse relation of correspondence which, even if it existed, could never be ascertained. But it is precisely at this point that Peirce would object. For what Peirce insists upon, and Dewey ignores, is that the kind of behaviour which constitutes science would never be indulged in unless men assumed that correspondence between statements and the facts were their goal. Dewey makes much of Peirce's definition of truth in terms of that upon which competent investigators might ultimately agree, but he forgets entirely the fundamental point upon which Peirce insisted they all do agree.

Whatever Dewey may mean in claiming to be a "realist", he cannot mean either what Peirce meant or what James meant. The crucial difference lies in his rejection of metaphysics. This at once cancels out the possibility of agreement with James, who conceived of realism *as* a metaphysic. And, for a slightly different reason, it cancels out agreement with Peirce. For while Peirce insisted that realism was a methodological presumption and not itself a metaphysic, he also insisted that it was a methodological presumption which made metaphysical speculation unavoidable. For, as I have pointed out, once you allow the methodological presumption that there *is* a real, you cannot arbitrarily rule out of court all speculation concerning what its nature is. I would certainly not maintain that I have said anything which would demonstrate that Dewey's larger doctrine is untenable, but I confess (*a*) that I am unable to describe what Dewey means by "realism" beyond showing that he cannot mean what either Peirce or James meant, and (*b*) that I am unable to describe any meaning of "realism" which is wholly consistent with what I take to be Dewey's larger doctrine. Russell, who is far closer to Peirce than is Dewey, has pointed out that even the presumption that the behaviour of investigation is "real" is inconsistent with any sweeping renunciation of the correspondence theory of truth.

Realism as Epistemology and as Metaphysics

Thus far I have attempted to do two things : (1) to distinguish between three senses in which the word "pragmatism" has been used ; and (2) to indicate how each of these "pragmatisms" has been associated with a different use of the word "realism". I turn now to some further uses of the term "realism" in America's recent past.

Let us return for a moment to James's brainchild—"neutral monism".[1] The American neo-Realists, like James, equated realism with this special metaphysic. Their arguments for their position consisted in large part in attacks upon the generative account of sensation which had been for some centuries the cornerstone of idealism, and their defence in its stead of a selective or presentative account of sensation. And, according to the

[1] In calling neutral monism James's "brainchild" I do not wish to imply that he was the original inventor of the doctrine. Though James was certainly unaware of it, Hume said many things which imply a neutral monist position. Moreover, James *was* aware of Mach. It nonetheless remains true that the phrase "neutral monism" was James's, and that the American neo-Realists were his philosophical progeny.

neo-Realists, a selective account of sensation logically entails neutral monism. It is worthy of note that the strength of twentieth-century realism lay in an analytic attack upon idealist epistemology. But epistemology was at every point associated with metaphysics. The destruction of idealist epistemology was presumed to eradicate idealist metaphysics; and realism itself was proposed as a neutral monist metaphysic which the epistemological revolution was held to entail.

There were, I think, two main reasons why the constructive position presented by the neo-Realists failed to gain a permanent foothold in the United States. The first was the degree to which their presentative and selective account of sensation led them, when facing the phenomena of error and illusion, to violate common sense. They were forced, in effect, to over-populate the universe. To say that mirages, bent sticks, and mistakes in general are floating about when nobody is aware of them, may have seemed sensible speculation from the perspective of technical analysis; but from the perspective of plain common sense it looked downright absurd. And say what we will, philosophers (in English-speaking countries at any rate) are loath to feel that they have lost touch with common sense. The second reason for neo-Realism's failure was perhaps more crucial. For as it turned out they were forced to over-populate the universe, not only by way of violating common sense, but also by way of violating the then still hallowed principle of empiricism. For the distinction between veridical and non-veridical perceptions repeatedly forced them to an account of the order within the sensations contributory to veridical perception, which order could not be defined without appeal to "gap-filling" sensations not in fact had by anyone but merely presumed to be there for the having.

In any event, the difficulties encountered by the neo-Realists led another group of men (and it is interesting that both American experiments in realism were group undertakings) to reject the presentative or selective theory, return to the representative or generative theory, flatly assert a dichotomy between content of experience and independent real—in short, to return to a dualist metaphysic. Here, then, is the interesting point for our present purpose: The Critical Realists, while they rejected the neo-Realists' association of realism with a neutral monist metaphysic, did not return to Peirce's earlier idea that realism is only an epistemological principle and independent of metaphysics. Instead they insisted upon associating realism with an alternative metaphysic—*i.e.* the special metaphysic of dualism. This is

perhaps most clearly seen in the best book which the movement of Critical Realism produced, Arthur O. Lovejoy's *The Revolt Against Dualism*.

Now, I have hitherto been taking the liberty of using the words "epistemology" and "metaphysics" without defining them. It is, however, by this time clear that the central point of this paper concerns the question whether realism should be conceived (*a*) as being merely an epistemological premiss independent of metaphysical alternatives, or (*b*) as itself constituting or entailing a particular metaphysical point of view. As this is the central question emerging, it would seem desirable that some clear distinction between epistemology and metaphysics be drawn at once.

This, however, is itself precisely the difficulty.

Everyone would agree, I suppose, that epistemology is a matter of analysis—of our methods, of the knowing process in general, and of the criteria of meaningful discourse. Presumably everyone would also agree that metaphysics is in some sense or other a matter of synthesis—that is to say, of hazarding some consistently worked-out proposal concerning the nature of things.[1] Beyond this point, however, agreement is more difficult to find. To maintain that epistemological analysis can be indulged in independently of any commitment on metaphysical issues is to maintain that one can analyse the knowing process, the methods of intelligent inquiry, and the problem of meaning without committing oneself to any point of view (explicit or implicit)

[1] Even the words "analysis" and "synthesis" may give rise to objections. It has been suggested to me that "clarification" and "synopsis" would be more neutral terms. If one conceives of epistemology as clarification, and asks to what degree a "clarifier" must indulge in synopsis, the answer is obviously that it is a matter of degree. Such language, however, seems to me vague on two scores. First, if a professional scientist who indulges in clarification is to be called a philosopher (or epistemologist) it seems to me clear that we need a distinctive criterion (apart from the notion of "clarification") to define the philosopher in the more special sense commonly used in academic circles. Is "degree of synopsis" such a criterion? Secondly, what, after all, is meant by "synopsis"? If one compares the dictionary meanings of "synopsis" and "synthesis" it is difficult to see that the one word will avoid any problems which hover about the other. I certainly concur with the opinion that it is a matter of degree how synoptic clarification may be, but I submit that it is equally a matter of degree how far people admit that their analysis entails commitment to a systematic point of view. I do not see that any change in terminology at this point significantly alters the nature of the problem I am raising. *No* terminology seems to me wholly satisfactory in this case, for it is being used to draw a distinction which is (this is my whole point) confused and unclear.

concerning the nature of things in general. Insofar as vast quantities of recent philosophical writings devote themselves to the rejection of metaphysics altogether, and insofar as these same writings surely themselves constitute elaborate epistemological investigations in the sense here defined, this claim of independence must be involved. It is, however, not necessary to reject metaphysics in order to support the claim. Peirce was an incurable metaphysician, yet in the context of our discussion it is Peirce who most clearly presents the thesis that a methodological premiss (realism in his sense) is quite independent of one's metaphysics.

The trends which I have so far been analysing seem on the other hand to provide fuel for the fires of those who contend that epistemology can in fact not be divorced from metaphysics. In James, in neo-Realism and in Critical Realism it is virtually impossible to separate the epistemological meat from the metaphysical bones—or, if you prefer, the epistemological bones from the metaphysical meat. Can you, as a matter of fact, go even so far as Peirce, and only so far (can you say that experimentation presupposes the independence of events upon observers, that conformity of fact to generalised prediction implies the independent status of concepts) without *eo ipso* committing yourself to some latent proposal concerning the nature of things ?

In recent years the positivists have devoted much energy to clarifying this problem and working out a solution. To the type of solution they offer I now turn. Once again, however, I should like to begin with some preliminary general and historical considerations which will have considerable bearing on what has now become my major theme.

The Relation of Positivism to Pragmatism and Realism

Positivism has been an extremely fluid movement. I recall that when I was an undergraduate a teacher of mine took great delight in ridiculing the positivists for constantly altering their opinions. I also recall, however, that the intended ridicule misfired. A good many of us felt that there was something open-minded and strong about a set of ideas which was developing and fluid, something which by contrast made this teacher's entrenched opinions (which happened to be idealism) seem closed-minded and weak.

Yet the fluid character of recent analytic tendencies makes brief and generalised discussion of them exceedingly difficult. I remind you that I am here using the word " positivism " in a very

broad and loose sense. It covers a great variety of highly divergent tendencies which often share little in common beyond the insistence that the task of philosophy is primarily one of analysis, and that the problems with which philosophers have traditionally been concerned must submit to the bar of analysis before continuing to flaunt themselves as genuine problems.

One more or less minor distinction between two avenues traversed by the analytic movement may be cited because of its help in marking the general temper of the movement in the United States. One form that analysis has taken involves an attitude of militant suspicion toward all technical language. The starting point for analysis is ordinary (usually English) popular idiom. Technical terms are considered genuine and treated with respect if and only if analysis of ordinary idiom reveals suitable justification for the first introduction of these technical terms. In all other cases technical language is suspect and problems which arise in such language are " pseudo ".[1] On the other hand, analysis has tended to take a quite different form when it starts from analysis of technical languages rather than ordinary (English) idiom. And it is in this latter form that analysis has tended to spread throughout the United States. In this form, philosophy starts not with ordinary idiom, which is itself in these circles the object of militant suspicion, but with the procedures and results of the more advanced sciences such as mathematics, astronomy and physics.

I wish, however, to emphasise a quite different distinction, a distinction which overlaps the one just cited. I shall assume that we are all familiar with the now classic division of discourse into (1) tautologies and analytic statements (*i.e.* statements which reduce to tautologies given appropriate definitions), (2) empirical hypotheses, and (3) the so-called " emotive uses " of language. I submit that the various forms of positivism have diverged along three distinct lines according as attempt has been made to place the propositions of philosophy itself in this classification.

The first (the *enfant terrible*) line was to hold that all respectable discourse is either of type (1) or of type (2) ; that certain propositions uttered by philosophers are of those types ; that all

[1] It is usually pointed out that the technical terms of science (often conceded blanket approval) are not subject to the criticisms levied against technical terms in philosophy. But this idea that there is a sharp distinction between the function of technical terms in science and in philosophy generally comes as a gratuitous and naive assumption rather than as an established conclusion. Those who are themselves immersed in the problems of theoretical science tend to be more sceptical of their technical terms than are a great many linguistic philosophers.

others, since they must be of type (3) are disreputable and should be discarded.

The second form, which began to emerge as everyone realised that the so-called " emotive uses " of discourse are not so black as originally portrayed, was to conceive of the task of philosophical language as that of decision and exhortation under the now more respectable type (3). It was recognised that few if any of the propositions either within traditional metaphysics or within contemporary analysis itself are tautologies. The propositions which constitute philosophical analyses are not themselves analytic propositions. Nor, indeed, are they empirical hypotheses. You cannot collect masses of evidence to make them more or less probable. Philosophers are more concerned with decisions and with the cultivation of attitudes than with either definitions or evidence. Understandably enough, some positivists themselves began to classify the traditional philosophical problems as " pragmatics ".

And there is yet a third form, which can best be described by contrast with the second. For it is held by some that it is a mistake to conceive that philosophy itself should do the deciding and the exhorting. The boundaries of the philosopher's task are drawn at that point where decision must be made and exhortation indulged in. Philosophy's task is to show that a man must make decisions, it is to reveal the difference between exhortation and proof, and clarify the legitimate scope and expectations of each. A distinction similar to the traditional distinction between ethical theory and ethical casuistry is applied to the entire range of philosophical discussion. Philosophy must not itself attempt to usurp the functions of ministers, judges and fathers. It has performed its legitimate function when it has clarified what can and what can not be expected from ministers, judges and fathers. One no longer encounters the claim (so it seems to me) that the propositions of philosophy reduce to tautologies and empirical hypotheses. But it is at the same time denied that those propositions are disguised explosions and exhortations. The classic tri-partite division of language would seem here to break down. For there is one use of discourse—precisely philosophy itself—which lies outside the list. It is that use of discourse which says when a proposition is a tautology or an empirical hypothesis or an explosion, and recommends rules for the legitimate functioning of each. Philosophy here seems to be searching for something like rules—unavoidable rules of legitimate procedure. And, in America at any rate, there is considerable respect being shown to the ancient metaphysical problems.

Logicians, in particular, are vitally interested in nominalism versus essentialism, and while they wish to distinguish between *e.g.* methodological essentialism and metaphysical essentialism, the latter notion is considered just as sensible as the former.

This paper began with a distinction between three senses of the word "pragmatism". It has now come to a distinction between three forms of the positivist or analytic movement. I find myself possessed with the idea that these two sets of distinctions parallel each other.

It seems to me that the "*enfant terrible*" stage of positivism committed the same error, though in reverse, which James committed—of attempting to tie together too closely the problems of meaning and of truth. James held that the conditions for determining meaning provided a key to the notion of truth; the early positivists held that conditions for determining truth provided the sole criterion of meaning. Each was wrong. James failed to see that there can be unpractical truths; the early positivists failed to see that unverifiable uses of discourse (*e.g.* commands, exhortations, and rules) can have clear meaning subject to definite rational controls.

Consider, too, the parallel between Dewey and the positivist tendency to conceive of philosophy as pragmatics. Dewey's proposal has been from the outset that philosophy should cease to have too high a conception of its precision and exactitude; should recognise that it is above all an "act of evaluation" (Dewey's key phrase); should recognise that evaluation is not abstract theory but a matter of concrete decision and exhortation in concrete problem situations. Considered as a technical analyst Dewey is desperately sloppy and unprecise. Yet many so-called positivists, having passed through the fires of long and rigorous analysis, have arrived at the same conclusion.

It seems to me that the more recent tendencies among the analysts are leading in the direction of Peirce, though admittedly there are few who would go all the way with Peirce. The tendency is to deny with Peirce that the pragmatic approach is sufficient for philosophy, and yet to recognise that analysis of scientific method is aimed at something deeper than either definitions or additional empirical hypotheses. (One might cite, for example, Karl Popper's lengthy and illuminating footnotes to *The Open Society and its Enemies*, especially Volume Two). Still and all, Peirce's own idea that what one is after are "necessary presuppositions" continues to have a too-Kantian ring. The phrase which I used a few paragraphs back, "un-

avoidable rules of legitimate procedure", was I confess my own, and would perhaps not be accepted. It is after all virtually synonymous with "necessary presuppositions". As for the logician's willingness to accept metaphysical problems as both sensible and important, this surely manifests a thoroughly Peircean attitude.

I am now prepared to return to the question of realism. Our brief glance at twentieth-century movements in American philosophy had raised the issue whether there is any connexion between realism as an epistemological or methodological assumption and realism in any of its metaphysical forms. More broadly, we were faced with the question whether analysis in general may be held to presuppose or conceal commitment to points of view concerning the nature of the world.

Now, as regards America at any rate, I do not think there can be any doubt that nearly all positivists are realists in Peirce's methodological sense. They simply assume that there are events and occasions independent of discourse and of the users of discourse. And the correspondence theory of truth is virtually unopposed. This is obvious in the favourite definition of truth: *i.e.* "*p*" is true if and only if *p*. Yet, though everyone in the most lively philosophical group in America is a realist, the issues discussed by neo-Realism and Critical Realism are all but dead. This is on the face of it an odd situation. Yet underneath it is simply based on the demand that a methodological presumption be separated from any special metaphysical pleading. So far, Peirce has emerged completely victorious.

But, as I have explained, Peirce held that epistemological realism though not itself a metaphysic makes metaphysics inevitable. Let us consider this side of the matter.

There is one point upon which I believe all branches of positivism would now agree. All would, I believe, confess that analysis does commit one to a "point of view". But all would agree to this only on the provision that what is meant by a "point of view" is neither a tautology nor an empirical hypothesis. They would disagree from that point on, but up to that point they would agree. The *enfant terrible* group would say that "points of view" are not the business of philosophy but of poets and preachers. The "pragmatics" group would say that "points of view" are the business of philosophers, but only insofar as philosophers themselves must indulge in persuasion and exhortation. The "Kantian" group (if I may with tongue in cheek so call them) would hold that "points of view" must be recommended as sets of basic rules.

I believe that this last group has yet to state its position clearly. I find that, although I heartily disagree with the first group, I can understand the claim that metaphysics should be eliminated from philosophical discussions. Though I disagree with the second group (disagreeing with less vigour) I can understand too the contention that metaphysics should be accepted as persuasion and treated as such. But the third group (with which I am most inclined to agree) seems to me to have been as yet unclear as to precisely what its stand is. I shall try to exemplify what I mean by this.

Consider the following passage from Quine's *Mathematical Logic*, p. 121: (This passage comes as an essential part of Quine's introduction of ϵ.)

> Discourse in general, mathematical and otherwise, involves continual reference to abstract entities. . . . One may prefer to regard abstractions as fictions or manners of speaking; one may hope to find a method whereby all ostensible reference to abstract entities can be explained as mere shorthand for a more basic idiom involving reference only to concrete objects (*in some sense or other*). Such a nominalistic program presents extreme difficulty, if much of standard mathematics and natural science is to be really analyzed and reduced rather than merely repudiated; however, it is not known to be impossible. If a nominalistic theory of this sort should be achieved, we may gladly accept it as the *theoretical underpinning* of our present ostensible reference to so-called abstract entities; meanwhile, however, *we have no choice but* to admit these abstract entities as part of our ultimate subject matter.

(I have selected this passage in part because it demonstrates that in the midst of his Platonic period Quine was quite willing, on certain conditions, to entertain a nominalistic interpretation. This indicates that much of the talk about his later "complete about face" in shifting from Platonism to nominalism is loose talk. Presumably he later became less doubtful about the possibility of satisfying the conditions laid down. But. . . .)

I would lay particular stress upon the three phrases which I have placed in italics. The first two: "in some sense or other" applied to the (for Quine) absolutely essential notion of "concrete objects", and the unusual phrase "theoretical underpinning" are remarkably vague phrases. What does Quine mean by "theoretical underpinning"? Analysts in general tend to treat logic as the quintessence of precision. They talk about logic as though it could never involve such a monstrously sloppy notion as "theoretical underpinning". But if you delve into actual logical systems you find this sort of thing recurring with astonishing frequency.

The third phrase to which I call attention is interesting for another reason. "We have no choice but" is the kind of language which implies the notion of necessary presupposition. It is the sort of thing that Peirce insisted philosophy is primarily designed to discuss. And in this Peirce was agreeing with Kant. (I add that I have not chosen Quine as a "bad" example. He is surely one of the best, if not the best, of living logicians. And the passage I have quoted is not unique; it is typical.)

In any event the implication of the passage—the implication either of the Platonism at the time of writing or of the nominalism of the envisaged shift—is that methodological matters and matters of what have traditionally been known as metaphysics are all wrapped up together. But how they are wrapped up—in precisely what sense metaphysical matters constitute the "underpinnings" of methodological matters (or vice versa)—is never explicitly stated.

I cite as another example of the type of unclarity to which I am referring, footnote 38 to Chapter Eleven in Popper's *The Open Society:*

> I wish to emphasize that I speak here about nominalism versus essentialism in a purely methodological way. I do not take up any position towards the metaphysical problem of universals, *i.e.* toward the metaphysical problem of nominalism versus essentialism (a term which I suggest to use instead of the traditional term "realism"); and I certainly do not advocate a metaphysical nominalism, although I advocate a methodological nominalism.

Popper, here and elsewhere, certainly makes it clear that he wishes to distinguish between methodological and metaphysical matters; he makes clear, too, that he considers metaphysical matters both significant and important; but I do not find that he ever makes clear what the *relation* between the two is. It is worthy of note that this footnote, which sharply distinguishes the two and insists that he is talking only about methodological nominalism is a footnote to a text in which several of Aristotle's metaphysical conceptions are under attack!

I believe that these examples are typical of the kind of remark one constantly encounters in present day analysis. I am unable to combat the impression that very basic questions, though recognised, are being gingerly sidestepped.

My own conclusion is perhaps evident. I do not follow along with those who insist on the rejection of metaphysics. I do believe that analysis is of the essence of philosophy and that one should try insofar as possible to keep it free at least from unconscious metaphysical presuppositions. But I think that in the

end it can never be kept wholly distinct. This means in effect that so long as we are all realists we are simply being lazy in refusing to meet such problems as Peirce and neo-Realism and Critical Realism and, for that matter, Kant bequeathed us. We are frightened by those recondite issues. But it does not make our position any stronger simply to refuse to discuss them.

Princeton University.

V.—SEMANTIC ANALYSIS AND PSYCHO-PHYSICAL DUALISM

By Arthur Pap.

In his highly significant and provocative book, *The Concept of Mind*, Gilbert Ryle undertakes to show that the Cartesian theory of two worlds, the physical world characterised by publicity and the mental world characterised by privacy, inaccessibility to all but one, is a "myth" created by misunderstandings of language. His method is an excellent example of the Wittgensteinian method of diagnosing the origin of puzzling philosophical theories as pointless, confusing departures from ordinary language. If, for example, a man puzzles how on earth it is possible ever to verify a proposition about the future since, after all, one cannot observe an event that has not yet occurred, his puzzle is of the kind which can be effectively dissolved by the Wittgensteinian treatment. The purpose of the following discussion is not a comprehensive critical review of Ryle's book. It is rather to show that his crucial arguments against the "privacy theory" of sensations—which is the aspect of psycho-physical dualism most relevant to epistemology—fail to establish what they are supposed to establish; and that they fail mostly because they exhibit a kind of exploitation of *ordinary usage* for purposes of criticism of philosophical theories which is characteristic of the Wittgensteinian method and which I consider to be an unwholesome feature of an otherwise most sanitary movement of semantic hygiene.

Ryle's basic thesis is that the theory of *mental acts* like believing, knowing, aspiring, results from the failure to see that sentences containing such psychological verbs are statements about dispositions, not about events or processes. I do not wish to add here to the literature pro and con the theory of mental acts; for, as stated, my specific purpose is to examine Ryle's arguments, set forth in chapter 7 ("Sensation and Observation"), against the Sense-Datum theory of Descartes, Locke and Hume, which survives vigorously in the epistemological writings of Russell, Broad and Moore. I cannot refrain, however, from expressing the opinion that Ryle's discussion is inconclusive because he has failed to formulate and discuss the basic semantic questions at issue between behaviourism and

dualism: (1) Are any behaviouristic sentences analytically entailed by such "mentalistic" sentences as "I remember to have seen this person before", *i.e.* are there any behaviouristic sentences with which a sentence of the latter type is logically incompatible; or is the relation between the mentalistic and the behaviouristic language only that of reducibility in Carnap's sense? If the latter is the case, as maintained by the contemporary form of *physicalism*, the most enlightened semanticist can go on believing in "Cartesian ghosts" since dualism is perfectly compatible with the public *confirmability* of mentalistic statements. (2) Is the above formal clarification of the issue really so clarifying, considering that the distinction between L-implication and F-implication is clearly defined for formalised language-systems but not for natural languages? If this scepticism with regard to the very tools of non-formal semantic analysis, which is becoming more and more vociferous, should be unanswerable, it would be anything but clear what Ryle is saying when he says that statements which the dualists interpret as referring to "ghostly" mental acts *are really about* behavioural events or behavioural dispositions.

I think the following is a fair summary of Ryle's basic criticism of the Sense-Datum Theory. Sensations are wrongly regarded as ways of observing, as though to see, to hear, to touch, to smell, were necessarily to observe something. The making of observations involves, indeed, the having of sensations, but it involves, moreover, an active element, the attempt to find something out, and above all, *what* one can properly be said to observe (where proper usage = ordinary usage) is always a *physical* object, event or process—where "physical" is used broadly so as to cover anything capable of being publicly witnessed, including, *e.g.* purposive behaviour. "... the ordinary use of verbs like 'observe', 'espy', 'peer at' and so on is in just such contexts as 'observe a robin', 'espy a ladybird' and 'peer at a book'" (p. 224). Ryle thus implies that the verb "to observe" (or synonyms like "to watch") is misused if it is used to describe an apparition in a dream or a state of hallucination—a contention surely open to debate though it will not be debated now. We may, I think, formulate Ryle's central criticism of dualistic epistemology as follows: Knowledge is a state of mind (= behavioural disposition?) terminating a successful inquiry; where there is no attempt to find something out, there is no occasion for knowledge to eventuate (this, of course, is again a statement about proper usage). It follows that the mere having of sensations does not constitute knowledge and that to speak of absolutely

certain, immediate knowledge, expressed by such sentences as "I see red", "I taste a bitter taste", is to misuse the word "knowledge" (notice that hyper-semanticist Ryle and anti-semanticist Dewey shake hands here). The Sense-Datum Theory postulates a curious kind of performance, called "pure sensation", which differs from observation ordinarily so-called in that it does not logically involve physical objects. But since an act requires an object (at least if it is expressed by a transitive verb, as are the verbs "seeing", "hearing", "touching", etc., in ordinary usage), and pure sensations by contrast to perceptions have no physical objects, the theory invents an appropriate kind of object called "Sense-Datum". With the help of this device the epistemologist constructs a pure sense-datum language whose sentences have the same subject-verb-object structure as sentences describing perceptions of physical objects: "I taste a sour sense-datum" resembles "I taste a lemon", "I see a round sense-datum" resembles "I see a round penny", etc.

It must be conceded, I think, that Ryle has correctly diagnosed the linguistic origin of one form of the Sense-Datum Theory, *viz.* the form which splits pure sensations into an act and a non-physical object called variously "sense-datum" or "sensum" or "sensibile". I am thinking particularly of Broad's theory of Sensa, which involves the postulation of non-physical entities which, like Platonic essences (except that they are "sensible" rather than "intelligible") are somehow there to become objects of the kind of awareness called "pure sensation": "It is quite certain that there is a difference between the two propositions: 'This is a red round patch in a visual field' and 'This red round patch in a visual field is intuitively apprehended by so-and-so'. Even if as a matter of fact there are no such objects which are not intuitively apprehended by someone, it seems to me perfectly certain that it is *logically* possible that there might have been" (Broad, *The Mind and its Place in Nature*, pp. 209-210). If all that Ryle had intended to establish by revealing that sensation is not a form of observation had been that it is misleading to speak of sensations as (mental) acts, I would have no further quarrel with him—except remarking what is perhaps obvious, *viz.* that one can consistently repudiate the act-theory of sensation and still believe in other kinds of mental acts, such as deciding, inferring, etc., as events defying definitional reduction to publicly observable actions. But evidently this accomplishment would not be enough to lay the Cartesian ghosts of traditional dualistic epistemology. Ryle claims moreover to dispose of the myth of *private facts* (or events), called "sensations",

which constitute the subject-matter of immediate, absolutely certain knowledge. But if it is linguistic confusion to suppose, with Broad, that there are mysterious non-physical objects, called "sensa", ready to be sensed, it does not in the least follow that it is linguistic confusion to suppose that there are events, called "sensations", which are not publicly observable in the way physical events are publicly observable. How does Ryle propose to show that it is nonsense to speak of sensations as events whose occurrence can be said to be *known*, yet never *publicly known*, *i.e.* known by anybody except the person owning the sensation? As far as I am aware, he uses two main arguments, which may be called: (1) the argument from the impossibility of a *pure sense-datum language*, (2) the argument from the impropriety of speaking of *observing sensations*.

(1) Before entering into explicit discussion of the first argument it might be pointed out that by no means all philosophers who speak the language of sense-data are guilty, like Broad and like Moore at the time of the "Refutation of Idealism", of the crime commendably lamented by Ryle, to split sensation-events into act and non-physical object. In Ayer's sense-datum language, for example, "I see a red sense-datum" is simply an artificial way of saying "I seem to see something red", the latter sentence involving the ordinary *perceptual* use of "to see", *i.e.* the use in which "I see x" entails that x is a physical object ready to be seen. (Strictly speaking, what is here called the perceptual use is not *the* ordinary usage but at best the most frequent usage. It is proper to speak of what one saw in a dream. If the perceptual use were the only proper use, a man describing his dream in the "I saw . . ." idiom rather than the "I seemed to see . . ." idiom would be speaking either improperly or falsely—which I doubt). A philosopher who finds it convenient to speak of sense-data, therefore, is not necessarily committed to the analysis of perceptual situations into a relative product involving over and above the observer and the physical object an intermediate entity: (x perceives y) = ($\exists z$) [(x directly senses z) and $z\mathrm{R}y$], where the interpretation of 'R' is the big headache for this kind of Sense-Datum Theory. But assuming that Ryle would grant this point, we may formulate argument (1) as follows. Let "S_p" denote a sentence reporting a perception, like "I see a robin", a sentence which could not be true, that is, unless a physical fact existed ("there was a robin at that time and at that place"); and let "S_s" denote a sentence which is just like an S_p except that it has no physical implication—understanding "implication" here in the sense of "analytic implication",

since by virtue of psycho-physical laws (correlations of sensations with brain-events) any S_s may *factually* imply physical facts. Ryle's argument then is that no S_s can be constructed, but that the moment one attempts to construct such a sentence, it inevitably turns into an S_p. "I see blue", *e.g.* can only be explained by translation into "I see the colour usually presented by such objects as A, B, C", and thus implies by its very meaning the physical proposition that there are (in the tenseless sense of "there are") objects of kind, A, B, C. Now, it is nonsensical to assert the occurrence of events (events described by S_s sentences) which it is in principle impossible to describe: "Whereof one cannot speak, thereof one must be silent" (the concluding sentence of Wittgenstein's *Tractatus* which latter, by the author's own ironical confession, is one big sin against this precept).

Now, I shall argue that what Ryle has shown is that the concept of an S_s *sentence with a communicable meaning* is self-contradictory —and in that sense meaningless—but has not shown what needs to be shown in order to refute the theory of private sensations, *viz.* that the supposition of the occurrence of pure sensations unconnected with physical situations is meaningless. Suppose I taste an apple-like taste (I mean this sentence in a way in which it does not *logically* imply the possession of taste-organs). The only way I can explain the meaning of "apple-like taste" to another person is by directing him to perform the physical operations which I believe, on the basis of analogical reasoning, will produce a similar taste-sensation in him. If the occurrence of taste-sensations were entirely uncorrelated with such operations as biting into an apple and thus were entirely unpredictable, we simply would lack the means of developing an inter-subjective language of tastes. But the very fact that this statement concerning the necessary condition of an intersubjective language of sensations is intelligible proves that the hypothesis of a world in which sensations are entirely uncorrelated with definite physical situations, a world in which one would never be in a position to say "it is probable that he is experiencing such and such a sensation now", is meaningful (*cf.* on this point, Schlick's article "On the relation between physical and psychological concepts", especially section VII, reprinted in Feigl and Sellars, *Readings in Philosophical Analysis*). That ordinary language contains no "neat" sensation-vocabulary, *i.e.* a vocabulary usable for constructing sensation-reports wholly devoid of physical implications, simply proves that ordinary language is inter-subjective. Naturally, there is no other way of communicating to a child the meanings of such words as "sour"

and "red" but to direct him to perform the kind of physical operations (tasting lemon juice, looking at blood) which one expects, be such expectations logically justifiable or not, will cause him to experience the kind of sensations designated by the words.[1]

Perhaps the following *simile* will clarify my argument. Suppose there existed a few circles to which I alone had privileged access, and that in my soliloquies I occasionally talked about those circles, using the word "circle". Assuming that no other circles exist and that I alone could see the circles that exist, how could I explain to other people what I mean by "circle"? Well, I could give a so-called *causal* definition by saying "a circle is the kind of figure which you would produce if you rotated a stretched thread kept fixed at one end". Suppose furthermore that a law holds in this hypothetical universe according to which a circle once produced by the described method is visible only to the person who produced it. It is clear that in this universe the meaning of "circle" could not be communicated by pointing to public circles but only by reference to straight-lines in rotational motion (and, incidentally, it would be theoretically impossible to make sure that one had communicated one's meaning, since, by hypothesis, one could not verify whether the experiment leads to the same result when it is performed by other people). Would it make sense for those people to speak of a possible universe in which there were circles but no straight-lines and no rotational motions? It surely would be a meaningful hypothesis, I should think, even though they would have to admit that in such a universe the word "circle" would lack a *communicable* meaning. I think that Ryle's argument from the impossibility of a neat sensation-vocabulary is parallel to the argument of that apostle of ordinary usage who got up and

[1] In the "Afterthought", at the end of Chapter 7, Ryle confesses to a bad conscience about having adopted the epistemologist's "sophisticated" use of the word "sensation". Ordinary language, he contends, contains no such expressions as "visual sensation" and "auditory sensation" (where ordinary language is contrasted with both philosophical and scientific language). I confess that this bad conscience impresses me as just another example of that pathological oversensitivity, cultivated in some circles of analytic philosophy, to philosophers' departures from ordinary language. What philosophers find epistemologically interesting about painful sensations (in one's leg or throat or eye), *viz.* the "privacy" of these events, is just a property which likewise characterises the events described by such sentences as "I seem to see a red patch". It is for this reason that they extend the term "sensation" to these events. Why should philosophers feel guilty about such departures from ordinary usage any more than scientists?

accused his speculating philosopher friend of contravening the ordinary usage of the word "circle" in speaking of circles isolated from straight-lines and rotational motions.

It might be added, before proceeding to examination of the second argument, that Ryle's arguments from ordinary usage against the epistemologist's language are largely irrelevant because they overlook that the epistemologist is engaged in what has been called *logical reconstruction* of empirical knowledge, and in this process inevitably and quite consciously develops an artificial language. No philosopher, to my knowledge, has ever claimed that atomic sentences, *i.e.* sentences which are logically simple in the sense that any two such sentences are logically consistent and logically independent, occur in ordinary language, nor that they describe isolated events—pure sensations actually occurring outside of perceptual contexts. If a sentence on the atomic level, like "I see red now", is to be construed as a sentence of ordinary English which may be true or false, it must be translated into "I see *something* red now" and thus ceases to be atomic: it has turned into an existential statement describing a perceptual situation [1] involving both sensation and a physical object. True, but irrelevant, for it is the epistemologist's purpose to reveal a level of language which is implicit in ordinary language though never spoken. A brief analogy may serve to remove the air of paradox from this concept of implicit language. It is impossible to imagine a sound-intensity divorced from any definite sound-pitch, and accordingly a sentence describing an experience of sound-intensity isolated from an experience of sound-pitch could hardly be found in ordinary descriptions of sound phenomena. Still, such sentences may have to be constructed if one wanted to make the complexity of the meaning of ordinary sound-language explicit ("a loud tone was heard" would not be sufficiently atomic, since "tone" already designates a complex of pitch, intensity and quality). As a matter of fact, it might be conjectured that some epistemologists have fallen into the confusions justly criticised by Ryle just because they did not go far enough in divorcing their artificial language of logical reconstruction from ordinary speech forms: "I taste sour" does not sound grammatical (unless it is a lemon that is speaking), so the epistemologist says "I taste a sour sense-

[1] My use of "perceptual situation" in this paper differs from Broad's usage in that hallucinations, *i.e.*, situations in which one seems to see an object which does not exist, are not perceptual situations at all according to my usage, while according to Broad they constitute a species of perceptual situation, called "totally delusive perceptual situations".

datum" in order to preserve the subject-verb-object form of ordinary sentences describing perceptual situations ("I taste a sour fruit"). And in some instances, as already admitted, these sentences of the epistemologist's own making have led to the mythical belief in sense-data as sensible, yet non-physical entities.

(2) Ryle undertakes to show not only that the mistaken assimilation of sensation to observation led semantically confused epistemologists to the invention of sense-data, but also that it is nonsense to think of sensations as possible *objects of observation*. "To observe a sensation", he holds, makes no more sense than "to spell a letter" (p. 206). If this is so, then the assertion that sensations, unlike physical events, cannot be observed by other minds is, indeed, of the same order as the assertion that letters cannot be spelled. The tacit implication of this comparison is evidently that the theory of absolutely certain knowledge of private sensations constituting the infallible basis of all our fallible beliefs about the physical world and other minds, vigorously defended again by Russell in *Human Knowledge*, is as devoid of significance as would be the emphatic assertion that only words, not their constituent letters, can be spelled. Now, Ryle's unquestionably superb knowledge of the English language denies me the right to throw doubt on his assertion that it is not proper usage to speak of "observing" or "witnessing" one's headache the way one may speak of "observing" or "witnessing" a car accident, while it is proper to use the verb "noticing" in this context (p. 206). But what of it? The psycho-physical dualist, I am sure, will gladly accommodate his language to the rules of good Oxford English and henceforth speak of sensations as events which cannot be *noticed* by more than one person at a time, unlike such physical events as car accidents. If Ryle allows the propriety of speaking of a "noticed headache", he should also allow the propriety of the sentence "I *know* that I have a headache—though you may doubt it"; but this is an excellent instance of private knowledge in the sense in which according to the dualist all knowledge of sensations is private.

But in fact Ryle's subtle investigation of the proper usage of such English verbs as "observing", "witnessing", "noticing", seems to me to be wholly irrelevant to the epistemological problem. The question is whether it makes sense to speak of sensations as events whose occurrence can be *known* with certainty by one person but not with the same doubt-precluding certainty by other people. If this question is answered affirma-

XUM

tively, the " Cartesian ghosts " have not been effectively banished. And the only relevant semantic question in this connection is whether it is proper to speak of " knowledge " of the occurrence of a sensation. We could easily imitate Ryle's method of observing carefully the speaking habits of educated English or American people, and cite in support of the propriety of such a usage of " knowledge " such sentences as " I know that these patches look all the same colour to me whatever they may look like to you—you can only know what they look like to you ". But it is more important to point out that Ryle is logically committed by his own statements to acceptance of the propriety of the expression " knowledge of sensations ". Observation of a robin, we are told, entails the having of sensations, though it involves more than that. Put in the formal mode, this means that the statement " I observed a robin " could not be true unless some statement reporting a sensation were true. But if it makes sense to speak of the truth of statements reporting sensations it must surely also make sense to speak of knowledge of the events reported.

I shall make some concluding remarks about the recent tendency of analytic philosophers practising therapeutic positivism,[1] a practice brilliantly exemplified by Ryle's *The Concept of Mind*, to regard the theory of the logical privacy of the mental as merely the upshot of arbitrary linguistic conventions. Consider the statement " it is logically impossible for one person to notice directly (notice that ' notice ' has been substituted for ' observe ') another person's headache, the way he could notice directly another person's sneeze; that's what makes the former event mental and the latter physical ". The usual reaction of therapeutic positivists to this assertion of the impossibility of direct knowledge about other minds is to ask: " Well, what would it be like to notice (directly) another person's headache ? Would this not amount to noticing a headache which is yours and not yours at the same time ? But then the event which you say is impossible is simply a self-contradiction, just like the event of spelling a letter (= decomposing into its constituent letters what is not composed of letters). Your talk about the ' logical privacy of the mental ', then, is as pointless as would be talk

[1] I am not using this label so as to imply the conception of philosophy as resembling closely *psychoanalysis* in method and purpose. By a therapeutic positivist I mean a philosopher who conceives of (good, worthwhile) philosophy as having primarily a therapeutic purpose, *viz.* the dissolution of typically " philosophical " perplexity by revealing the linguistic confusions which cause such perplexity.

about the impossibility of spelling letters". I would be the last one to deny that this kind of analysis of the expression "logical privacy of the mental" is immensely clarifying. It is important to realise that it would be self-contradictory, in terms of current meanings of the expressions "thought", "direct knowledge", to suppose that A could have direct knowledge of B's thoughts in the sense in which B could have such knowledge.

One important consequence of this semantic fact is that telepathy should not be interpreted as direct *knowledge* of another mind's mental states but rather as direct belief, *i.e.* belief not derived from physical observations, about another mind which stands a high chance of being correct. This may be demonstrated as follows: if we credited A with direct knowledge of B's present thought in the sense in which we credit B with such knowledge, then we would accept the following two premises as conclusive evidence for the proposition describing B's mental state: (1) A asserts that B is in this mental state, (2) A's assertion is honest, *i.e.* A believes what he says. But the fact is that while we would accept these premises as providing conclusive evidence if A and B were one and the same person (not that I mean to imply that we could have conclusive evidence for the truth of these premises in the sense of "conclusive evidence" in which the premises, if true, would be conclusive evidence for the conclusion), we don't if they are different persons: we would first ask B directly about his mental state before we accepted A's telepathic *belief* as an instance of *knowledge*. The point I am arguing can be made particularly evident if we suppose that the mental state of B which A is alleged to have direct knowledge of is a state of belief. For in that case the claim would either mean that A is introspecting his own belief which happens to be shared simultaneously by B, or that the state of belief is literally a common constituent, a point of intersection as it were, of the mental histories of A and B. In the first case there is no telepathic situation at all, in the second case we are violating the ordinary usage of the expression "mental state" in somewhat the same way as we would violate the ordinary usage of "particle" if we spoke of the possibility of the same particle occupying simultaneously different places.

Yet, I submit that after all the virtues of this kind of semantic analysis have been recognised, one is left with an enormous *non-sequitur* when one turns to the claim that the dualist's concept of the mental defined by logical privacy is *nonsense*. All that has been shown is that the assertion of the impossibility of certain, immediate knowledge of unowned mental states is an analytic

consequence of the meanings of the expressions "certain, immediate knowledge" and "unowned mental state", and not a factual thesis like the impossibility of catching a certain thief. The further premise which would have to be added and substantiated before the argument of Ryle and associates could be accepted, is that those meanings are gratuitously stipulated by dualists in sheer violation of ordinary usage. To illustrate such gratuitous deviationism in linguistic conduct: if a philosopher arbitrarily defines "there is evidence for *p*" to mean "*p* is logically demonstrable" and then announces that it is impossible to produce evidence for empirical propositions referring beyond present experience; or, if he defines "seeing a physical object" to mean "seeing simultaneously the whole surface of the object as well as all its past and future states" and then announces that common sense is mistaken in believing that physical objects could be seen; or if he defines "seeing the *same* object A again" to mean "seeing A again with exactly the same properties it had previously" and derives from his definition the paradoxical consequence that it is impossible to see the same object twice—in cases of this sort philosophers are rightly blamed by therapeutic positivists for building perplexing philosophical theories by means of nonsensical (*i.e.* self-contradictory) concepts misleadingly associated with meaningful everyday expressions. The question before us is: Is a philosopher similarly guilty of arbitrary departure from ordinary usage when he says "I can be *aware* only of *my* feelings, not of anyone else's; I cannot *know for certain*, though I may conjecture, what other people's thoughts are"?

That we do, in everyday discourse, speak of awareness and certain knowledge of other people's feelings and thoughts, is admitted. In fact, such sentences as "I am clearly aware of her jealousy; I know he thinks I do not know about his affair with Mrs. X" are, *in the sense in which they are intended*, frequently known to be true. The philosopher's claim that such sentences cannot be known to be true or to be false seems, therefore, like the same sort of wilful paradox, brought about by surreptitious departures from the ordinary meanings of words. Yet, the semantic situation here differs in an important respect from the kind of "philosophical perplexity" illustrated above. In the cases I adduced, the tricky philosopher convicts the man in the street of cockily holding unjustifiable beliefs, by substituting for his vague but applicable concepts, precise but self-contradictory (and so inapplicable) concepts. But this idle game must be distinguished from the fruitful game of revealing

ambiguities in ordinary language which are unnoticed by the man in the street. Suppose I asked the ordinary mortal for whom knowledge of other minds is simply a fact, no problem at all: "Are you aware of your friend's loneliness in the same sense of 'awareness' as you might be aware of such a feeling in yourself?" Surely, Mr. Common Sense need not be analytical above average to come to see, after some reflection, that there is a big difference between introspective awareness of a feeling, which is in a sense not easy to analyse infallible, and inference to similar feelings in others, which is in an obvious sense fallible. The essential point is that this concept of direct, introspective knowledge of a mental event is a concept in actual use, not a concept wilfully constructed by philosophers intent on belittling common sense. The dualist's statement "it is impossible to know that another person feels lonely the way one can know the occurrence of such a feeling in oneself" is, indeed, analytic of the relevant meaning of "knowledge". But the statement's analyticity would convict it of triviality only if the concept analysed were one artificially constructed and never used by non-philosophers.

If the advocates of common speech whose attack on psychophysical dualism I have criticised were to use their method consistently, they would, I suspect, approach practically any philosophical analysis of a concept in use with the same cynical attitude. Take, *e.g.*, Hume's denial of a necessary connexion between cause and effect. This denial, of course, is exactly the same kind of "violation of ordinary usage" as the dualist's denial of the possibility of knowing for certain what another person is thinking or sensing or feeling: we do say, quite often, "this effect must necessarily follow . . . ", and in the sense in which this kind of statement is intended we frequently have excellent grounds for asserting it. But the ordinary man will quickly cease to be puzzled by Hume's statement if the latter is clarified in the usual fashion, *viz.* as the denial of the logical demonstrability of causal laws. The fact that the ordinary man now discovers that he never really held the belief just demolished by the subtle philosopher does not convict the philosopher of having launched a Quixotic attack on wind mills. For, as Moore has taught us, the purpose of analytic philosophy is not to *refute* common sense beliefs nor to declare them as unfounded, but to *clarify* them. Thus Hume, whatever his own confusions may have been, has the merit of having clarified the difference between inductive and deductive inference, empirical and logical connexion, a difference obscured by the

XUM

ambiguous use of the words "reason", "ground", "consequence" (even "cause", in the Aristotelian tradition) in both senses. The concept of the mental as the logically private, in contrast to what is publicly observable, is similarly the result of analysis of confused meanings. And I submit that if by psycho-physical dualism be meant no more than the conception of the mental here defended against Ryle's ingenious assault, there is no reason to suspect psycho-physical dualists of metaphysical backwardness isolated from the progressive tools of modern semantic analysis.

University of Oregon.

VI.—COUNTERFACTUAL CONDITIONALS

By Robert Brown and John Watling

We have an impressive number of uses for counterfactual conditional statements, and to some recent writers it has seemed that the analysis of these into synonymous indicative statements is a problem of considerable importance. Professor Chisholm, for instance, wrote that "the philosophical problems which this question involves are fundamental to metaphysics, epistemology, and the general philosophy of science".[1] However, both he and Professor Goodman,[2] after detailed examination, conclude that the problem is still unsolved. Is it?

Consider the peculiarities of the project: an attempt to replace a counterfactual conditional of the form 'if x were ϕ, then x would be ψ' by an indicative statement having the same meaning. Since the rather complicated formulae which can and have been produced for this purpose do not work, we may mistakenly accept Professor Chisholm's alternatives. We may, that is, either tinker with a formula for eliminating counterfactual statements in favour of indicative ones, or believe that counterfactuals, and with them subjunctive conditionals, refer to some kind of connexion between relations, qualities, events and things. Such "irreducible connexions" are a vested interest with which we can dispense, and the formulae are irrelevant as well as unsuccessful. We suggest that the reason why both of these assertions are true is that the problem, so-called, is misconceived; we cannot eliminate counterfactuals, nor should we want to do so.

I

Briefly, the analysis of counterfactuals should parallel that of "fugitive propositions". Somewhat as Mr. Nowell-Smith has shown [3] that tensed sentences refer implicitly to a context which cannot, without paradox, be made explicit in the sentences themselves, so it can be shown that the use of counterfactuals presupposes contextual sentences that are neither implied by the

[1] "The Contrary-to-Fact Conditional" in *Readings in Philosophical Analysis*, New York, 1949.

[2] "The Problem of Counterfactual Conditionals", *Journal of Philosophy*, p. 44, 1947.

[3] *Analysis*, vol. x, no. 5.

counterfactuals nor part of their meaning. This is most clearly indicated if we recall the way in which we know whether a singular subjunctive conditional is being used counterfactually or not. It is perfectly correct to say 'If this bear were a Kodiak, it would be brown', when as a matter of fact we don't know if it is a Kodiak. The additional information that it is not a Kodiak must be given to us before we can know that we are making a counterfactual statement. Thus singular counterfactuals cannot be of the form '$\sim \phi(a) \,.\, \phi(a) \supset \psi(a)$' because '$\sim \phi(a)$' represents the information supplied by the context. For the same reason the fact that the counterfactuals are not directly testable is hardly puzzling. It would be puzzling only if counterfactuals implied the falsity of their antecedents instead of presupposing that falsity. It has been suggested that "A man who says 'If this bird were a raven, it would be black' thinks and thinks rightly, that he is saying more than 'This bird is not a raven'."[1] But the man is not *saying* the latter at all, he is presupposing it. There is no sinister problem arising from the circumstance that false antecedents can never be realized and tested, for while we cannot directly test counterfactuals we can supply evidence for them and give rational grounds for their assertion, namely, universal statements and initial condition statements. Some cases are quite simple. The singular subjunctive conditional used by Mr. Pears[2] "If this bird were a raven, it would be black" holds, and its contrary 'Even if this bird were a raven, it would not be black' does not hold, when the statements made by the following sentences hold: (*a*) 'Any raven is a bird' (*b*) 'Any bird which is a raven is black' (*c*) 'This is a bird'. When we are challenged on the subjunctive conditional just quoted, all we can do is to produce statements like those made by (*a*), (*b*) and (*c*), claiming that the first is analytic and the latter two are well confirmed. We can point out that if this bird happens to be a raven, so that (*c*) is fulfilled, then (*b*) entails that it will be black; but neither the counterfactual nor its contrary implies that the bird will be or will not be a raven. That extra piece of information must be given to us.

Most cases are more complicated. These include the counterfactuals used in history, law and social science, ones like 'If the Labour Party had called a general election in 1949 it would have lost'. The difficulties of supplying acceptable evidence here are obvious. There are fewer relevant, well tested universal and initial condition statements available for this example than there were for the previous example. From (*a*), (*b*) and (*c*) together

[1] *Analysis*, vol. 10, no. 6. [2] *Analysis*, vol. 10, no. 3.

we can derive the singular material conditional 'If this bird is a raven, it will be black'. Since it does not have the same truth value as a counterfactual they cannot be equivalent but the additional knowledge (*d*) 'This bird is a raven' allows us to infer the testable prediction sentence (*e*) 'This bird will be black'. In defending a counterfactual we appeal to the statements corresponding to (*a*), (*b*), (*c*), (*d*) and (*e*) as evidence. We say, 'It so happens that $\sim$ (*d*) is the case, yet if (*d*) were the case, then (*d*) in conjunction with (*a*), (*b*) and (*c*) would entail (*e*)'. If this evidence were not acceptable, either because our universal and initial condition statements were not well confirmed or else because of a logical error, then short of abandoning the counterfactual, we would have to produce other evidence and derivations. But what kind of evidence can we supply in the example of the Labour Party? By assembling a long list of trend sentences, inexact functions and general or universal sentences, along with a vast number of initial conditions sentences, we might be able to show that relatively to these our counterfactual has a certain amount of confirmation, although the evidence would be precarious and liable to challenge on several grounds, including that of ignoring relevant conditions. This is the situation with regard to counterfactuals used in economics and politics, as well as with those applied to people's characters and motives. In point of fact, arguments about motive explanations are often disagreements over the choice and independent testability of the motive statements serving as initial conditions in the explanation.

The amount of evidence we can produce varies considerably. In some instances the counterfactual is entailed by its evidence statements. A play in a game, for example, can so occur that it is both the last permissible play and the one upon which the gain or loss of the game depends. Assume that it is a penalty kick made by the Bruins at their opponents' goal. After the play is completed and the Bruins have won, we can say counterfactually, 'If the Bruins had missed their penalty kick, they would have lost the game'. This statement is jointly entailed by certain rules of the game and the statement describing the circumstances of the last play. At the other extreme there are counterfactuals for which no evidence can be produced because of the conditions under which they are made; yet we still want to say that such statements are either true or false. Consider the counterfactual 'If I had thrown the dice once more I would have won'. Confirming evidence for statements like this is often completely lacking. The 'Why?' question is answered by 'I

had a hunch' or 'I just know I would' where these remarks are not replaceable to the speaker's satisfaction by any description of chances, of relevant laws or special conditions operating. It is logically precluded that any evidence *pro* or *con* can be given for the results of dice-throwing are due to chance, and hunches about single throws are attempts to predict or retrodict when no grounds for rational belief are obtainable. By education, of course, we might be able to eliminate the use of counterfactuals under these circumstances.[1]

Various misguided efforts have been made so to interpret counterfactuals that they will imply the statements serving as their evidence, and then to substitute for counterfactuals a synonymous or equivalent set of indicative statements as desired. Once the possibility and desirability of the programme are accepted, obstacles to its fulfilment multiply generously. The course of the predicament runs like this. First, due to the important uses we have for counterfactuals, we want to say that much of our knowledge is of what would occur if conditions were other than they actually are. We then go on to ask quite naturally, 'What kind of unusual knowledge is this? Can knowledge of what does not occur or did not occur be expressed in indicative statements? What is the relation between antecedent and consequent in a counterfactual statement? Do counterfactuals give us untestable knowledge?' These questions create a puzzle which becomes rapidly inflated; until it is solved we do not seem able to provide an adequate interpretation of natural laws, and without this nothing can be done in the philosophy of science.

Professors Chisholm and Goodman realized, of course, that in counterfactuals the consequents follow from the antecedents only when the antecedents are conjoined with certain other relevant statements which hold in fact. They wanted to be able to characterize these relevant statements in such a way that the description could be incorporated into an indicative formula that would enable us to distinguish the conditions under which counterfactuals held. It is not necessary to discuss the inadequacies of the attempt, for no attempt need be made. The distinction desired is obtained by an appeal to evidence, and our evidence for asserting a counterfactual cannot be implied by that counterfactual. A counterfactual *refers* to its evidence statements in the sense that unless they are true or to some extent confirmed we cannot defend our counterfactual and so do not employ it unless we are ignorant, careless or lying. To have

[1] We owe most of this paragraph to Professor A. J. Ayer.

counterfactual knowledge is simply to have knowledge obtained by interpolation and extrapolation on the basis of natural laws. The "irreducible connexion" presupposed by counterfactuals is that asserted by natural laws. In asking what kind of connexion holds between antecedent and consequent in a counterfactual we are asking about natural law connexions without intending to, and we are assuming also that the relation of antecedent and consequent is something like material implication or entailment, but yet perversely different.

II

Is it correct then to argue that all material implications are irrelevant to counterfactuals ? This cannot be the case, of course, if natural laws are equivalent to or synonymous with universal material implications, for these laws are presupposed as evidence in our use of counterfactuals and are hardly irrelevant. Are natural law statements equivalent to (or synonymous with) universal material implications ? Neither entail singular counterfactual conditionals ; both entail singular subjunctive conditionals. The latter also follow from universal subjunctive conditionals and these, in turn, are equivalent to universal material implications.

It is clear that not all universal material implications are universal scientific laws. For instance when the Beast says to Beauty " You will never see me again, except every evening at 7 o'clock " his announcement can be written as a universal material implication but hardly counts as a universal law. Certain restrictions must be put upon the kind of statements occurring as antecedent and consequent before a universal material implication can be said to be a general law. But without considering what these restrictions must be it is still possible to argue that all scientific laws are expressible as material implications. Whether this will turn out to be an interesting assertion will depend upon the interest of the arguments that may be given for or against it.

Against the argument that laws can be expressed as material implications it is said that such expressions might be true only by accident. That is, perhaps the real law would differentiate between several sets of conditions for the fulfilment of the antecedent, whilst the material implication is true because in fact, and by accident, only one set of these conditions ever turns up. Someone living in the tropics might formulate the law ' Everyone likes ice cream ' and this be true by accident in the tropics

because the temperature there is always high. And if high temperatures prevailed all over the world all of the time then the accidental character of this law might never be realized. And the philosopher arguing against the material interpretation might say, 'It is the connexion between hot weather and the liking for ice cream that is important and this connexion is not stated in a law that may be true only by accident'. The most extreme type of this argument is found in examples such as 'All four legged swans are black' which are true as material implications because there are no four legged swans. Once more the argument is that it is only accident that makes these laws true; no connexion between the antecedent and the consequent is asserted.

Another way of stating the same argument is to maintain that accidentally true material implications do not allow the assertion of counterfactual conditionals. 'If this man were living in a cold climate he would like ice cream' or '. . . not like ice cream', is, according to this argument not a legitimate deduction from 'All men like ice cream' because 'All men like ice cream' is only accidentally true. Yet another way of expressing the argument is to object to the confirmation of a law by instances where the antecedent is false.

A similar problem arises in the analysis of counterfactual conditionals themselves. Any analysis which makes a counterfactual equivalent to the logical product of a singular or universal conditional and the negation of the antecedent of this conditional allows it to be verified by merely verifying the falsity of the antecedent.

Mr. Kneale arguing against the equivalence of natural laws and material implications gives an interpretation of "(x) . raven $(x) \supset$ black (x)". The statement for which this is the formula he says, is equivalent to the statements made by the following three sentences: "(1) There has never been an observed raven that was not black; (2) There has never been an unobserved raven that was not black; (3) There will never be a raven that is not black." Only the statement made by (1) is confirmed by the evidence; (2) and (3) are not, but only they are useful to us in our inference from the observed to the unobserved.

Let us consider the argument from accident. A law states the conditions which are causally necessary for the truth of the consequent. Variations in any of the factors mentioned in the antecedent will affect those mentioned in the consequent and no variations in any other factors will do this. Only then can the law be used for prediction.

A law that ignores one factor may be said to be true by accident if it happens that this factor never varies. If this factor does vary the law may be falsified and we might say that previously it had only been true by accident. A law may be said to be true by accident if we can identify a factor which was not mentioned in the law and whose constancy was not questioned when the law was applied successfully. But though the new law is better than the old, still we cannot be sure that circumstances will not turn up revealing the necessity for extending the antecedent even further. For there is no *a priori* limit to the number of factors that may causally interact. So that there always remains the possibility that subsequent investigation will reveal other factors upon whose variation the event depends, and therefore that previous confirmations of the law may be said to be accidental. But how are we, given two laws both of which have never been discredited by any observations, to distinguish one as accidental and the other as not ? We might decide by the amount and variety of confirmation that the law has received, or by consideration of other laws with which these two are compatible or incompatible. In Mr. Kneale's example of ' All ravens are black ' this is said to be accidental because if ravens had lived in polar regions they might well have become white and then the law would be untrue. That is, there is an appeal to another law, that all birds living in snowy lands are white, in order to suggest that ' All ravens are black ' is accidental, not a law at all. ' Any four legged swan is black ' can be sensibly affirmed or denied; how much evidence we can produce one way or the other will depend upon how much knowledge we possess about swans, especially about the relations between their legs and their colour. Similarly, in the case of natural laws about freely falling bodies and frictionless surfaces, our evidence will consist of exemplified natural laws. In all cases then, the word ' accidental ' is used for recognizable characteristics of universal material implications ; either they appeared true for a while and were then falsified or there is good reason to believe, because of other laws or insufficient confirmation, that conditions for their falsification might arise.

There are circumstances in which we can recognize a law as accidental but none where we may be certain that it is not accidental. The parallel in this respect between ' accidental-law-like ' and' falsified-verified ' is exact. Just as we have no use for ' completely verified ' so we have no use for ' certainly non-accidental ' ; but we do not withold the name of ' law ' from universal statements that are not completely verified, and

neither should we withold it from universal material implications that are misleadingly called 'accidental'.

The use of 'accidental' and 'non-accidental' does not depend on the truth of counterfactuals which refer to conditions that have never occurred, but on observations or on other established laws. Professor Popper has proposed that "it may be useful and, for some purposes, necessary to assume a general principle stating that every kind of event that is compatible with the accepted natural laws does in fact occur in some (finite) space region". The argument that there must be no instances of vacuous truth if universal material implications are to count as laws runs like this: If there were possible values of the antecedent that did not occur, then it is possible that if they had occurred the event referred to in the consequent might not have occurred. As long as vacuous cases of the implication remain there is still a chance that the implication is only accidentally true. But in the first place this proposal would be insufficient to achieve the required result. A law could only be said to be non-accidental if it could be shown to hold no matter what the state of all other factors might be, including those which, though observable, remain to be discovered. This would require an infinity of worlds. In the second place we have seen that there is a perfectly good sense of 'accidental' which does not need such absurd conditions for its satisfaction.

Consider Boyle's law: by it we can calculate from the volume of a gas at a certain temperature to its pressure at that temperature. Interpreting it as a material implication allows us to do this for any volume that we may find experimentally; it also allows us to state counterfactuals about volumes we have not measured, for these become vacuously true when the antecedent is false and ordinary factual statements when the antecedent is fulfilled. The statement of universal material implication does not symbolize an intrinsic connexion between the volume of a gas and its pressure. But as there could not, logically, be any observation of such a necessary connexion or causal relation it is no demerit of universal material implications that they do not symbolize one.

Having examined this asserted equivalence, consider the question of whether those material implications which are not laws are irrelevant to counterfactuals. It is true that universal material implications do not imply or entail singular counterfactuals, because the latter are not of the form

$$'-\phi(a)\,.\,\phi(a) \supset \psi(a)',$$

though they are only used when ' $-\phi(a)$ ' represents what is believed to be the case or holds in fact. But it is not true to say that the universal material implication " There has never been a raven that was not black and there will never be a raven that is not black " is irrelevant to the counterfactual " If this bird were a raven, it would be black ". It is argued that " There has never been a raven that was not black " is compatible with the statement that if ravens had lived in a snowy environment their offspring would have been white. It is, the argument continues, only an historical accident that they have not lived in such regions, just as it is merely an accident that there have never been any non-black ravens. Knowing, as in fact we do not, that there never has been such ravens would not permit us to claim " If some inhabitants of snowy regions were ravens, they would be black ".[1]

This argument owes what plausibility it has to the example chosen. Nevertheless it is unsound. A counterfactual cannot, strictly speaking, be incompatible with a universal material implication, but two evidence statements for a counterfactual can be contradictories. Yet the absence of incompatibility is not relevant in this case. The material implication in question is relevant in that it constitutes part of the evidence upon which the counterfactual is asserted. The material implication does not so commit us that if our evidence for making it were to be different (*e.g.* arctic ravens) from what it is, we should still make it. With different evidence we might well frame different material implications.

" If some inhabitants of snowy regions were ravens, they would be black " is supported by referring to the universal material implication " There never has been a raven that was not black and there will never be a raven that is not black ". What gives credence to the belief that the second statement is not evidence for the first ? Only that we are reminded of another material implication about protective colouration in animals, and if this were so it might be something like ' There never has been a raven whose colour was not matched to its environment '. Black arctic ravens would provide a counter example to this, just as white arctic ravens would to ' all ravens are black '. Depending upon which of these universal material implications we believe the better confirmed, we shall accept the one or the other, and with each its respective counterfactual, " If some inhabitants of snowy regions were ravens they would be black " and " If

[1] *Analysis*, vol. x, no. 6.

some inhabitants of snowy regions were ravens, they would be white". An appeal to the difference of evidence ought not to be mistaken, however, for the denial that a particular universal material implication is evidence for a particular counterfactual.

III

When the opponents of the thesis that laws are universal material implications state that counterfactuals could not be obtained from them they are wrong. They may be right in asserting that counterfactuals, analysed in the way they choose, cannot be obtained. They choose to analyse them as though the statements are not counterfactual but factual. It is not even 'If he hadn't crossed the Rubicon, then . . . , but he did'; rather it is 'If he hadn't crossed the Rubicon, then . . . , and he didn't'. Whether a law may be legitimately applied in such cases is scarcely worth arguing, for the application of the law will certainly not lead to any observable consequences. That a universal material implication does not provide such deductions is, therefore, no argument. It is not enough that Boyle's law has been confirmed and will be confirmed in the future; it is also required that it should be confirmed in all possible worlds.

When someone is faced with several courses of action the form of argument used will be a conditional, but not a counterfactual conditional. A scientific law in the form of a universal material implication will provide such conditionals, for we are only interested in the results of the argument when the antecedent is fulfilled. The view of counterfactuals that we have been arguing against amounts to carrying this procedure over to past events, ignoring the actual decision which has taken place, and considering the conditional obtained as verified true in some other sense than its determination by the actual truth values.

Suppose we say that the law is just a rule for inferring one sentence from another. The rule is used by substituting a determined sentence for the propositional function forming the antecedent and so obtaining a determined sentence for the propositional function forming the consequent. We might then say that the rule was no more than an apparatus for making these inferences and that apart from the possible substitutions it was meaningless.

Now suppose that instead of substituting determined sentences we substitute sentences which have not been determined empirically, giving us a scheme of a large number of possible

inferences. These possible inferences will yield inferences from one sentence to another when the sentences are tested and the truth values put in. This way of carrying out the substitution arrives at the same result as the first way but splits the process up into two stages. The interpretation of counterfactuals as independent of material implications amounts to stopping at the first stage of this substitution process. But at this stage the usage of the symbols still only provides rules into which substitutions may be made, and these are just as meaningless apart from the substitutions as the original material implications between propositional functions. To suppose that laws in some way describe a world in which all these inapplicable inferences can be made, and that because universal material implications do not do this they fail as laws, is obviously incorrect. It is precisely the same fallacy as to suppose that laws describe necessary relations between cause and effect. Of course there is no objection to entertaining a sentence which has not been tested and making deductions from this provided that it is realised that the application of these calculations requires sentences that have been decided by observation.

In supposing that a law is a rule of inference it is necessary to allow it a truth value, so as to distinguish those rules that are to be applied from those that are not. And the decision of this truth value requires the processes of confirmation and disconfirmation. By virtue of this truth value the rule of inference has a descriptive use. Those counterfactuals which are asserted quite independently of any evidence, as in hunches about dice throws, are statements analogous to the results of the first stage of substitution in a law ; but the law from which they might be obtained has no degree of confirmation.

On this interpretation conditionals, even when particular but undecided sentences have been substituted, provide only schemata from which inferences may be constructed. Counterfactuals, derived from conditionals whose antecedents have already been decided as false, are inference schemata that will never have any application, but which are not thereby meaningless. For we could recognise situations in which the schema would apply, although we know empirically that such situations will not arise.

The main points we have tried to make are these : that the analysis of counterfactuals must be such that it is physically impossible, or believed impossible, to decide them ; that the evidence for counterfactuals is evidence for general laws and the applicability of these laws, but that a counterfactual is not a statement of a law or of initial conditions. Neither is it a

conditional obtained from a law by substitution, nor the conjunction of such a conditional with the denial of its antecedent, for both of these would be decidable at once. Rather it states the validity of an inference, usually derived from a law, that can for practical reasons never be made. Since there are no facts to decide these inferences no general law need describe them.

University of London.

VII.—FREE-CHOICE AND THE DEMANDS OF MORALS

By Frank B. Ebersole

I should like to re-examine one rather limited but, I think, significant aspect of the multiform, confused, old "free-will" question. It is argued sometimes that an action must be uncaused before it can be praise or blameworthy. More commonly it is maintained that praise, blame, responsibility and punishment positively require that all actions must be caused. I believe I can show that both these views are mistaken. These wrong conclusions seem to be drawn primarily because the importance of the term "person" in judgments of praise or blame is not fully appreciated. This is surprising since all theorists recognize that praise and blame are bestowed only upon persons; and the conditions under which a person may be held responsible or punished are generally recognized as the cardinal point in question. My present purpose is not the ambitious one of attempting yet another analysis of 'person' or 'self-identity', but I believe I can show that the usual moral arguments for either indeterminism or determinism are based on an incorrect or meaningless conception of selfhood. Whether incorrect or meaningless it is not possible to say, for some conception of 'self' is presumed in these arguments but it is not explicitly acknowledged or stated. I suggest that the mechanist metaphors have got the better of many philosophers when they discuss free will, even philosophers who, in contexts which call for explicit discussion of 'self', know better. Like the legendary capitalist who controls the whole town but has never been seen, an invisible "self" is held responsible for crimes and rewarded for meritorious acts. Whether or not this is so, at least for some reason the time of a blamed action in relation to the time of blaming is seldom given due attention.

There are four different kinds of moral question whose answers provoke discussion of freedom: (I) Ought (or ought not) that person to have done what he did? Or, ought (ought not) that person to have done what he did not do? (II) Is that person to blame? Or, is that person to be praised? (III) Is that person responsible (usually, but not always, for a crime)? (IV) Should that person be punished (or rewarded)?

Any statement which can be construed as an affirmative answer to questions of type (I) I shall call judgments of obligability. When an affirmative answer is correct or justified, I shall call the action which would be cited to justify its correctness an obligable action. Affirmative answers to type (II) questions, I shall call judgments of praise or blame, and actions which would justify such judgments, praiseworthy or blameworthy. Affirmative answers to type (III) questions, I shall call judgments of responsibility. Only persons, never actions, are said to be responsible. Still a judgment of responsibility could be justified only by citing an action or actions for which a person is responsible. Affirmative answers to questions of type (IV), I shall call judgments of desert. As with 'responsible', only persons are subject to punishment and reward, but they are punished or rewarded for actions. The punishment or reward could not be justified without considering certain actions which the person had done.

I use the term 'judgment' here to denote a class of utterances or marks which have in common (in addition to certain similarities) either a designation or a certain regular effect on behaviour. I shall not try to decide to what extent such moral judgments designate and to what extent they act as non-designative causes of behaviour and attitude. However, when we look into the grounds or justification for their use, it is their efficacy as changers of attitude and behaviour which must occupy our attention. Consequently, I should like to add to this list of judgments a fifth class of occurrences which is not a judgment but a kind of moral action, *viz.* (V) punishment or reward. Such actions share with the four preceding judgments a certain distinctive kind of effect on attitudes and behaviour. The conditions under which a person is justly punished or rewarded prove to be similar to those under which a person is correctly held obligable or held to blame.

Each of the five classes contains positive and negative expressions. What I shall have to say will apply, with minor alterations to positive expressions, but for the sake of brevity and simplicity I shall confine what I say to the negative. Negative judgments of each type I shall call judgments of condemnation. What I shall say, then, applies to judgments of condemnation and acts of punishment. I shall call the sum of these two classes, condemnations. To this class belongs utterances of the specified forms which use the terms 'blame', 'punishment' and, when the action for which the person is held responsible is a wrong action, judgments using the term 'responsible'. To this class also belong judgments which use the term 'ought' and 'ought not' in the

following two ways: (*a*) "You did *x* and you ought to have done *y*", and (*b*) "You did *x* and you ought not to have done *x*". Lastly, the class of condemnations includes the executions of punishment.

A complete condemnation always specifies an action. This is plain with judgments using the terms 'ought' and 'ought not'. With blame and desert, judgments are complete only when the action is specified for which a person is to blame or for which he is responsible. Likewise a punishment is incomplete unless it is for a specific kind of action.

Now we often ask, "Why is that person condemned?" A correct answer to this question seems to require these three parts: (1) The person was free (or the action was free, or the choice to do the action was free), (2) The action "belongs to the person"; "it is an expression of his character", and (3) the kind of action he did is wrong. I mean to use the term 'wrong' here as it is often used of classes of actions such as in "It is wrong to be dishonest". In this sense 'wrong' seems to be equivalent to 'ought not' in a certain sense in which 'ought not' is predicated of the same class of actions. "It is wrong to be dishonest" in most contexts seems equivalent to "One ought not to be dishonest". In this sense, to say "One ought not to be selfish" is perfectly compatible with saying of a particular action, "It is not the case that he ought not to have acted selfishly". It would be compatible if the person in question was not free to act unselfishly. But the sense in which 'ought not' is applied to classes of actions is not the same sense as 'ought not' in judgments of obligability. 'Ought not' or 'right' as applied to classes of actions does not "entail can". 'Ought not' as applied to particular actions does "entail can", and it is with this "entailment" that I shall be concerned. It seems likely that either use of 'ought' can be defined contextually in terms of the other, perhaps thus: One ought to do *x* (*x* is right). ≡ If he is free and has no superior obligation, then one ought to do *x*. On the left hand side 'ought' does not "entail" freedom; on the right hand side 'ought' is used in the sense it has in judgments of obligability, and does "entail" freedom.

Part (1), the attribution of freedom to persons or actions or choices, is of concern to morals because even if we knew (*a*) what 'right' and 'wrong' mean, and (*b*) what kinds of actions are right and wrong, and (*c*) what counts as a person, and (*d*) which actions are expressions of a personality; we would still be unable to justify moral use of our knowledge unless we also knew (*e*) the meaning of 'free' and (*f*) which actions are free. In what

follows I intend to discuss what we know about (*e*) and (*f*), and what we need to know to justify condemnations. A discussion of these points positively requires some discussion of (*c*) and (*d*), although this is not always recognized.

What is the Question?

Any abstract question usually begins by being so poorly formulated that one generally has to examine the answers given before he knows at all accurately what is asked. The free-will question, I think, is no exception. Even among those whose philosophical dispositions and methods are similar there is wide disagreement in answers given. This disagreement is, in a sense apparent only, for I believe it indicates a difference in question asked. For example Broad (*Determinism, Indeterminism, and Libertarianism*) examines the meaning of 'ought' and 'ought not' and concludes that in the moral usage of these terms they may be correctly applied only to actions, specifically actions such that a person doing such things either has a very low ideal of human nature or a very weak and unstable desire to approximate the ideal he has. Also, more significantly, the terms 'ought' and 'ought not' apply only to actions such that the person doing them could have had higher ideals or could have willed more strongly or persistently to live up to his ideals (pp. 24-25). The meaning of 'could' in these phrases, Broad insists, means more than 'would have if'.

On further examination Broad concludes that the additional component in the meaning of the term 'could' is (1) the action or volition must have one among its "causal ancestors which is not completely determined by other events", and (2) it must be a desire (pp. 35-37). Broad calls such events 'accidental', and thinks that judgments of obligability based on them are mistaken (p. 42).

John Wisdom (*Problems of Mind and Matter*) gives an analysis of 'blame' in which he concludes that it "entails ultimate self-direction", *i.e.*, among the causal antecedents of a blameworthy action there must be at least one event concerned with will and which is not completely determined. Thus both Broad and Wisdom think that the notions of obligability and blame respectively "entail" indeterminism.

On the other hand, the far more usual analysis finds the meanings of 'punishability' and 'responsibility' not only consistent with complete determinism but requiring it. Schlick (*Problems of Ethics*) reaches this conclusion about 'responsibility'.

Baylis[1] analyses both 'obligability' and 'responsibility', and concludes that both require determinism.

Why this apparent difference in conclusion? I think it is because different questions are being answered. Four different questions are usually left undistinguished: (A) What is the correct analysis of the terms in question as they are used in the rough and tumble of moral affairs? (B) Under what conditions do we in fact make judgments of responsibility, blame, etc.? (C) Can we justify such judgments? and (D) What justifiable judgments can we make? Questions (C) and (D) are distinguished in case we find the usual judgments unjustifiable. If it should turn out that judgments of censure and praise are wrong, the conditions under which we make judgments applying such terms as responsibility, blame, etc., would not help us in reaching definitions. Therefore question (A) must be distinguished from (B). With these distinctions in mind, I think it becomes easy to see that there are two closely related reasons why these different answers are reached.

(I) I think that of the four questions enumerated, Schlick and Baylis, *e.g.*, have question (D) in mind. They try to state the conditions under which we can justifiably impute responsibility and execute punishment, and so are led to offer definitions of 'freedom', 'responsibility', 'punishment', etc., which are applicable. Schlick actually says he is answering question (A): ". . . there is required only an analysis of the concept, the careful determination of the meaning which is in fact joined to the words 'responsibility' and 'freedom' as they are actually used". (Strangely enough, he then goes on to show that ordinary usage confuses two senses of 'law' and two senses of 'determinism'.) Broad and Wisdom on the other hand begin with question (A), an attempt to analyse what they believe to be the "common-sense" meanings of obligability and blame. Then they turn to questions (B) and (C). Broad finds judgments of obligability unjustifiable. Wisdom finds that if they are correct they entail some unlikely form of pre-existence of the individual before his birth.

(II) Schlick and Baylis concentrate at crucial points in their discussions on the conceptions of 'responsibility' and 'punishment'. Asking about the ordinary sense of freedom required for 'punishment' and 'responsibility' leads to different results from asking about the sense of freedom required for obligability. The reason for this, I believe, is that condemning to punishment

[1] Charles A. Baylis, "Rational Preference, Determinism, and Moral Obligation" (*Jour. of Phil.* vol. xlvii, No. 3, February 2, 1950), pp. 57-63.

is much more serious than making a mere statement of 'ought not', hence it has more generally provoked ethical thought about its purpose and justification. Whether or not I am always aware of it, when I ask myself what kind of freedom must characterize a man's wrong action before I will take part in punishing him, I am influenced by the possibility that my first hasty impulse may be unjustified. The consequences of punishment are grave and usually largely unknown. My judgments here are more cautious, and when called for, more apt to reflect past examination. In short they are morally more enlightened, and this enlightenment has almost become part of "common-sense".

However, this does not obliterate the distinction between questions (A), (B) and (D), nor minimize the importance of making the distinction. In fact there is too often confusion between (A) and (B), *i.e.*, between analysing the ordinary, conversational meanings of words and analysing common-sense beliefs expressed with the help of these words. I should like to mark this distinction hereafter by referring to ordinary, conversational, or every-day *meaning* or *usage*, and confine the term 'common-sense' to *beliefs*. Strictly speaking we can ask for a justification only of sentences expressing beliefs. There is a derivative and different sense in which we can ask for a justification of meanings, which means merely "are the definitions which explicate ordinary meanings convenient or useful? *i.e.*, can we elegantly express our most defensible beliefs using these terms"?

I think it is fairly clear that as society discovers that there is no justification for punishment in a retributive sense and that judgments imputing responsibility and wreaking punishment in order to "give a man his due" are never justified, the word-form 'responsibility' comes to change its meaning. What might have meant retributive punishment comes to mean reformatory punishment. If unenlightened commonsense demands retributive punishment it demands that certain things hold about the cause of an action. Non-retributive punishment demands that something be true about the effects of an action. In the former case the causes must have been part of a person's character, in the latter case the effects must become part of a person's character. In either case the action must be free, but only in the former does 'free' refer to something about the causes of the action.

Ordinary Meanings of 'Free'

The meaning of 'free' where attributed to an action by moral judgment is both vague and ambiguous. Either analysis, (1)

a form of indeterminism, or (2) a kind of determinism, seems correct depending on whether we begin with ' freedom ' demanded for obligability or ' freedom ' demanded for punishment. Putting aside the question of justification for the present, I will attempt an explication of what common-sense seems to demand for obligability—and show why it is inapplicable.

G. E. Moore's analysis provides a good starting point. Moore suggested first in *Ethics* that it might be a sufficient condition for an action to be free [1] that the agent whose action it was would not have done it if he had chosen not to do it (pp. 13-16). This proposal conforms to an important condition for the ordinary use of ' free action ' and one on which many analyses go wrong, *viz.* the condition that an action does not have to result from a conscious choice in order to be called free. Later, after considering the criticism of A. C. Garnett,[2] Moore was led to think that this analysis is a gross departure from everyday usage, and that in addition to the condition that an action be such that a person would not have done it if he had chosen not to, it must be such that the person could have chosen not to do it. This surely is what ordinary usage demands. Yet as Moore had previously pointed out in *Ethics* this result is puzzling. In seeking to clarify the phrase ' could have done something else ', we expand ' could ' to ' would have—if he had chosen ' and then add ' he could have chosen '. The word ' could ' returns to be analysed. Now either it means the same thing in its second occurrence as it does in the first or it means something else. If it means the same thing, then ' could choose ' means ' would choose if he had chosen to choose '. It has been urged against this possibility that we never have the experience of choosing to choose.[3] I believe this is wrong. The experience is common enough, *e.g.* I may be considering whether to go to the ball-game or read a book, and I cannot decide. Someone suggests that I decide by choosing a card from a deck. Let us say : red card, I go ; black one, I read. Now I must choose whether or not to choose. In order that the term ' choose ' should mean the same thing in both uses, I must, of course, deliberate as to whether or not it is fair to choose in this fashion. It is still true that moral decisions do not usually involve choosing to choose, and it is clear that if this

[1] I have deliberately substituted ' free ' for ' voluntary '.

[2] A. C. Garnett, " Moore's Theory of Moral Freedom ", *The Philosophy of G. E. Moore*, Library of Living Philosophers, Northwestern University, 1942, pp. 179-199.

[3] D. J. O'Connor, " Is there a Problem About Free Will " ? (*Proc. Arist. Soc.*, 1948), pp. 33-46.

analysis does not conform to the ordinary meaning of the original 'could', it will not conform to the second, or to any other. The second occurrence of 'could' must mean something different. It is better to substitute back our original word 'free' for the second meaning of could. Thus we get 'free to choose'.

This much has been achieved: ordinary usage does not demand that a free action be the result of conscious choice as long as it could have been. An action is called free only because it could have been chosen: thus free action may be defined as one that results from free choice. The question then concerns certain conditions antecedent to an event which was an act of choosing or which could have been an act of choosing but need not have been. Also it seems clear that the crucial preceding condition need not be one which immediately precedes the possible act of choice. It is common knowledge that the way a person acts in the present is determined in part by past choices he has made. If a person has many times throughout the academic year chosen to read science fiction rather than study philosophy, it becomes increasingly hard for him to prepare for his final examinations. He will read science fiction instead, and we say he ought not to do this. Similarly a person might choose to smoke opium instead of studying. It becomes increasingly unlikely that after many such choices he shall be able to choose not to smoke opium at examination time. Now, however, we feel 'he ought not to have chosen' is not applicable at examination time. Still he ought not to have chosen in the beginning. Sometimes freedom is an hereditary property, sometimes not. How does ordinary usage differentiate these two examples?

The desires which are built up for opium are so great that there has been available to no known person a counter-desire sufficient to overcome it. Or else they are qualitatively new desires with which the person has had no more experience than a year old child with his overwhelming desire to throw everything he can get his hands on. A strong counter-desire to science-fiction, however, could throw the balance. With sufficient exertion a person could choose to study in preference to reading science-fiction. This much is clear; but of more importance for the present problem is another condition. The desires at the very moment of choice determine the choice, but prior to that time in some cases the desires are not themselves determined. A knowledge of the desires just prior to the choice, together with the laws of nature would enable a Laplacian psychologist to predict the choice, but a knowledge of all previous characteristics of the person without the last desires would not be sufficient

for the prediction. This is the very simplest case. The desires may also be predictable. Then an action is not free unless somewhere among its causal conditions there are psychological causes of the desires which are chosen, and the desires determining this choice are not predictable. How far back this condition of unpredictability or indeterminism must be depends upon the specific detailed circumstances which common-sense knowledge does not possess. Hence ordinary usage is vague. But one is foisting off on the common-man something he will not accept unless a 'free-act' contain causes somewhere in the history of the person which are undetermined. More particularly, there must be somewhere in the person's past at least partially uncaused desires which have previously been subject to choice and which result in a decision which makes up part of a cause which causes, etc., an obligable act of choice.

There is an ambiguity in the ordinary usage of the terms 'desire' and 'want' which is worth noting, because much bad psychology stems from a failure to observe it. In one sense of the term desire, to say that I desire something is contradicted by the statement that I do not know what I desire. For example, when I say I want (or desire) chocolate cake, it may be correct to conclude, after getting chocolate cake and hence fulfilling my desire, that I did not need chocolate cake. It would never be correct to conclude I did not want (or desire) chocolate cake. To say 'I desire x', in this sense is to say 'I am thinking of x, imagining my having and enjoying x', etc. In this usage of desire, it does not make sense to say, 'I thought I desired x', in the same way we say 'I thought I was awake'. In this sense desire denotes a conscious state.

There is another sense of 'desire' and 'want' in which it is correct to say 'I did not know what I desired'. Take, *e.g.*, the sentences, 'I always wanted (desired) my own pleasure but I did not know it' or 'I desired nothing but applause when I gave my speech, but I thought I desired a job well-done'. Here 'desire' does not include thought. It does not denote a conscious state. It does not, in fact, denote a state at all, but a motive, an impulse or tendency to act in certain ways under certain conditions. It is this second sense of 'desire' I have used in what I believe to be a description of the common-sense view of choice. A set of final desires, together with certain events determines a given action. The desires are subject to change. Persons do not always have the same desires, nor do they have desires in the same degree of intensity. Among the causes believed to alter the strength of desires (in the sense of motives)

are desires (in the sense of conscious states). A desire whose intensity is continually lowered, is lost. An act of choice would seem to be an act of resolving a conflict among conscious desires which adds weight to a motive, and so resolves a conflict among motives. Some conscious desires coincide with a possible exhibition of a motive, and so strengthen it.

It would be wise, no doubt, to mark this ambiguity of 'desire' by using two terms. But I shall have no occasion hereafter to use 'desire' in any except the meaning 'propensity to certain actions'. I do not pretend to give an adequate description of an act of choice, and consequently I shall not have further occasion to refer to the conscious desires which clearly make up a part of the business of choosing. A rougher characterization of what I believe to be common-sense indeterminism will suffice.

To summarize, the characteristics which ordinary usage demands for a free action seem to be: (*a*) the action must be that of a person; (*b*) it must be the kind of action which could have been chosen, *i.e.*, the desires which, together with certain circumstances determined it, were sometime subject to choice; and hence, (*c*) somewhere in the causal ancestry of the action is a desire which is undetermined; and lastly, and quite obviously, (*d*) the action must not have been such as would have occurred even had it not been desired, or subjected to choice. This last condition is clear, because if the action must not be determined by the person's past, then more definitely, it must not be determined by events in which the person could have had no choice.

I think it quite certain that the ordinary intelligent and intelligible person who is uncontaminated by philosophy or science insists that something like condition (*d*) is part of his meaning of 'free' when it is used in moral contexts. If this condition ever holds in fact, then the "principle of determinism" or "the principle of causality" as it is sometimes (mistakenly, I believe) thought to be required for the procedure of science is false. This is the crucial condition for the rest of the discussion. I should like to state it more precisely, and in so doing eliminate the troublesome word 'cause'. This statement requires, as was promised, some clarification of condition (*a*).

Take an event for which a person would be blamed or held responsible, say, wilfully kicking a friendly dog. Assume, then, that '*b* kicked a dog at time *t*' is a true sentence. The flexibility of conversational language will not permit *b* to be hardened into a "logically proper name", for we can just as easily designate the same occurrence in terms of an event, *viz.* '*b*'s kicking a dog took place at time *t*'. The first sentence suggests the myth of

b, the pure ego. The event sentence has the advantage of lessening this suggestion, but is subject to its own temptations. The second sentence is one of a large set of sentences which completely describe the public-behaviour-biography of person *b* up to time *t*. The possessive '*of b*' or '*b*'s' becomes a property of the event and is apt to suggest a corresponding class of events, the mental events which constitute the inner experiences of *b*. Since this type of sentence emphasizes the temporal order of the public events, it suggests a temporally ordered second class of events which constitutes the "mind" of which the public behaviour events are expressions and effects. Such a "theory" of mind is generally devised out of unrequited love for simplicity. It would normally assign a few pervasive dispositional properties to these mental events on the model of unobservable physico-chemical events which are assigned only the general dispositional property of obeying the laws of physics.

In terms of this "theory" of mind condition (*d*), the condition of indeterminism would say that a physicist's complete description of the world in terms of unobservable events, and a complete psychologist's mental biography of *b* up to a certain time (the time of the undetermined desire) when conjoined with the laws of nature would not entail the sentence describing the event for which *b* is held responsible. (I assume that the laws of nature include statements correlating certain arrangements of microscopic and/or unobservable entities with the dispositional properties of observable things). But this will not do. This definition of indeterminism is quite compatible with what is meant in everyday language by every motive's being determined. That is, it is quite compatible with the statement that a complete description of the person together with the laws of nature do entail the sentence describing the guilty act. In ordinary language *b*'s kicking the dog would be accounted for in terms of certain irritating occurrences which happen to *b* and *b*'s sadistic predisposition and his tendency to brood over his irritations. It is his brooding and sadistic tendencies which are believed by common sense to be a little uncaused.

Only in the unlikely case that it should turn out that a description of the mental biography of a person together with the laws of psychology should entail a complete description of the person, would this definition of indeterminism contradict the ordinary notion of determinism. In particular it must be the case that statements about the mental events, together with laws, entail statements about a person's acquiring, having and losing wants and desires. Desires are not events any more than elasticity or

hardness are events. The presumption that a chemical theory of postulated entities like molecules would enable a chemist on the basis of observations of rubber and cement in a liquid state to predict that they would be respectively elastic and hard when solid has borne fruit for the development of chemistry. There is no indication that the analogous presumption is fruitful for psychology. Nor is there any use observing a person in anything analogous to a liquid state. Even if the analogous presumption were fruitful and applicable in psychology, such a guiding principle must not be enumerated among the laws of the science. We would never know whether such a definition of indeterminism as this theory states was even materially equivalent to the ordinary man's conception until actual detailed predictions were made in large numbers. Even so, we are not seeking material equivalence; we are seeking an analysis of the meaning of undetermined desire as it is used by an ordinary person in everyday moral practice.

In his ethical thinking it seems quite certain that the ordinary person is interested in uncaused motives, and in his moral judgments he is interested in altering motives. He is concerned with changing fairly permanent features of the character of persons who do wrong. When he affirms condition (*a*) of a 'free' act, he means that the action is the exhibition of a trait, a motive, a disposition. The condition of indeterminism, therefore, must read: a complete description of the world including the behaviour-biography and the character of a person up to a certain time (the time of the beginning of the undetermined motive), when conjoined with all the laws of nature, does not entail the event for which *b* is being condemned. The "principle of determinism", I take it, would say that a description of the world during a certain time span together with the laws of nature entails a complete description of all events. ("Softer" or emergent forms of determinism would insist that the description of the world must include a segment of the biography of *b*, or of some human ancestor of *b*, or of some animal ancestor of *b*—depending on where in the scale of nature the emergence is believed to be.)

However the condition that a free act proceed causally from an act of choice which is wholly or partly undetermined constitutes only a necessary not a sufficient condition for the ordinary usage of 'free'. The sufficient condition requires the "self" or "person" to choose. The ordinary understanding of a freely chosen act includes not only the condition that the choice be uncaused, but strangely, also includes the condition that the person himself "threw the balance" one way or the other in the act of choosing.

This " person " or " agent " is not constituted by a sequence of past events, and dispositions, for they stand only in a relation placing some limitation on the possible desires. Among the possibilities which remain open, however, there is no past event or events observable or inferrable which determine the choice which " throws the weight " to one or another. It is the self, the ego, the person acting presumably not through the past but directly on the undetermined desire. With these metaphors, " throwing the weight ", " acting directly ", the common-sense man completes his story. Philosophers take up his story at this point and construing the analogy with simple-minded ontology, infer, " the timeless subject of ideas ", " the transcendental self ", " the pure ego ", etc. I need not repeat here the reasons for thinking that the descriptions of such entities are meaningless.

What Kind of Word is ' Free ' ?

There is another condition for the ordinary use of ' free ' which is sometimes held. An examination of this condition reinforces the conclusion that the conversational term ' free ' is never applicable. The everyday notion of a free act not only includes the conditions that such an act must result from a choice which is not determined by past events but perhaps by a " pure ego ", it also includes the condition that free choices are data of experience. The ordinary usage of ' free choice ' is such that I believe I recognize a free choice when I experience one. I consider for example, two possible courses of action ; I think about the consequences. Here and now I picture the presumed results of each. My desires waver. Finally " I throw my weight " to one. I recall the means leading to that one. I will the actions necessary to reach it. I claim to recognize this wavering, considering, and final resolution as the free act.

Curiously, if certain circumstances have preceded the action and they are called to my attention, I revoke my judgment, " That was a free act ". Suppose I am peeved with my neighbour, and after much brooding I coolly walk into his house and pull the tail feathers out of his pet toucan. In retrospect I now recognize this as a free act. But suppose now that I had submitted to a hypnotic experiment the night before my evil action, and that under hypnosis it had been suggested to me that I would pull the tail feathers out of my neighbour's toucan the next afternoon. Suppose also that it had been suggested that in ill-humour I should hatch this particular scheme and deliberate before coolly executing it. My sinful deed which I had taken for a free act, I now agree was not free. I was mistaken, and precisely

because the antecedent conditions are such as to preclude free action.

It might be suggested that the mechanism of hypnosis does not affect the freedom of the action, but rather changes the person to such a degree that the action does not belong to him but to someone else whose deliberations he shared. It remains a free action but the free action of another person, who largely displaced the original person for a day or two. However, the evidence of memory and the continuity of feelings and attitudes would scarcely suggest, let alone substantiate, so severe an hypothesis. I think it may be concluded that 'free act' is not, as might have been expected, a term like 'ball' or 'tree' whose analysis consists in a semantical rule stating the conformity of 'ball' or 'tree' to nursery samples. This result could also be established simply by noting that the phrases used to describe a free act such as "picturing alternatives" and "throwing my weight" indicate efforts to infer arrangements among mental events on the analogy of physical ones.

Our preceding analysis shows, however, that 'free act' is a phrase whose analysis can be stated in a definition. The definition includes reference to conditions which may precede the action by a considerable time. Moreover, it is negative in form; it refers to the absence of a cause. How could we ever be directly aware of the absence of a strand in the causal net which preceded an action? The fact seems to be we could not, and again the conclusion seems inevitable that our judgments of obligability attributed to persons because specific actions are free are probably always wrong. At best if they are ever correct, they are correct entirely by accident.

It might be suggested that the word 'free' as ordinarily used in moral contexts is ambiguous, in one sense, being an analysable term like 'mermaid' which can be reduced to perceived things and in another sense entirely designating certain introspective data: a gestalt composed of deliberating, entertaining, weighing, etc. If 'free' is ever used in this latter sense in moral contexts, it cannot be an applicable sense. It would then denote moments of conscious choice. An action is condemned precisely because it is free. But it is certain that actions are condemned which are not free in this sense. This alone is not a good argument against this use of 'free' as the relevant one. Such actions may be wrongly condemned. But we do condemn them because they are free, and we examine before condemning to see whether or not they are free. We do, when we examine them, seek to discover not whether they were conscious choices, but whether the

person could have chosen if he had been conscious. A person who kills a child with his car while "unconsciously" staring out of the window would be condemned because his action was free. Ordinary discussions of its freedom would consider whether or not he could have been aware of the choices before him between applying the brake, swerving, or hitting the child, and whether the habit of "unconscious" staring could have been avoided by past choices. If he could have been aware he would ordinarily be said to be free. Actual awareness is not the condition relevant for the type of judgment I have been considering.

How are Condemnations Justified?

To ask for a justification of condemnations is to ask whether they fulfil the purpose for which they were used and whether the purpose is right. Justifying judgments of condemnation does not consist in deducing them from more general principles. Justifying condemnations is like justifying the laws of logic, not like justifying any particular argument. The justifications may be given either in terms of (I) a retributive, or (II) a non-retributive purpose.

(I) A condemnation is justified retributively when the person is caused to have pain and remorse. There is no further question. Anyone holding a retributive theory of condemnation is satisfied as soon as he is convinced that the guilty person has been pained and feels regret. Again to consider the meaning of these statements, we need to know what is meant by a person. Arguments for a retributive effect usually insure the identity of the person with the help of the "pure ego". But if it is the same identical pure ego, who is subject of change but never changes, which is to be caused pain and remorse, then condemnation might just as well occur before a crime as after. In fact it might just as well not occur at all, for the pure ego is insulated against the temporal nexus of causes and effects. Chastising a pure ego is either like teaching ethics to a quadratic equation or like teaching metaphysics to a bandersnatch.

On the other hand if a person be conceived as a complex string of mental events, then there is no difficulty about the temporal succession of condemnation and remorse. A person commits a crime; we discover the crime; we take action against the person. Then it is a temporal segment of the person's mental biography somewhat later than the criminal act which receives the results of the condemnation. The trouble with such a cinema-man is just that it would be impossible for him to feel remorse. For

feeling remorse is neither an event like feeling a momentary itch nor a string of such events. A person who has been feeling remorse since last Thursday cannot be described as having a continual series of feelings or pangs. A few pangs he must have had. But they must be tokens of a propensity, when idle or when tempted to repeat his wrong deed, to think of his crime and its motivations with regret. Feeling remorse is a disposition which may characterize a person for a few days or a few hours. It is exhibited in pangs of regret, but it tokens possible future pangs and types of overt behaviour. For moral condemnation, a person must be subject of propensities, traits and dispositions which may be acquired and lost. Their loss or acquisition must be caused by occurrences such as moral judgments, punishment or reward, either alone or together with other dispositions. This is required either for a retributive or non-retributive theory of condemnation.

A strictly retributive theory would have to argue that effects of a condemnation had nothing to do with their justice. No one would hold such an extreme position. In order to give a justification on retributive grounds some semblance of plausibility we must distinguish the immediate from the remote effects of a condemnation.

The immediate effects of being condemned, then, are pain and remorse. The remote effects are the reformatory effects of pain and remorse on the character and later activities of the person. A retributive theory, modified in this way, would then argue that the immediate effects were justified even if the remote effects were different. It is simply " just " that an evil person should suffer, *i.e.* that we should add suffering to the personal history which contains the record of a past wrong deed. That it is just to punish a wrong in this sense could be defended solely on the basis of its conformity to common-sense belief. I believe this is not the common-sense view of punishment. It may appear to be so because when asked to decide on it one cannot imagine the artificial circumstances required for evaluating a series of events in complete disregard of their likely consequences.

A defender of a retributive theory would have to hold that pain and suffering when added upon a crime were right regardless of the consequences. Suppose a man committed a crime, and we positively knew that although he was responsible (in the ordinary meaning) he would never repeat it. And also suppose we positively knew that punishing him would inflict great pain and produce bad consequences. The retributive theory would say these latter conditions had nothing to do with the case;

the punishment was right. This is surely in violation of the common-sense view of what is right.

It is also argued in favour of a retributive view of punishment, that a non-retributive theory does not account for the meaning of the terms "punishment" and "what he has coming to him". It is said that punishment does not mean reform, and "what he has coming to him" does not mean producing the best consequences. If the ordinary meaning of these terms is retributive, then they are never applicable because there are no 'free choices' in the ordinary sense of 'free'. My conclusion is that we cannot rely on common-sense beliefs. Therefore we can always ask for the justification of judgments of the kind I am examining and hence ask whether or not the terms 'ought', 'responsible', 'punishment', etc. are applicable. I am forced to conclude that not much can be said for a retributive justification of condemnation.

(II) A retributive justification claims that a person's state before condemnation simply merits condemnation. That is all there is to it. A non-retributive justification is based upon the person's state after condemnation. The reason for censure is that it influences the person who has done wrong in such a way that he will be less likely to do in the future that which occasions the censure.

Of the four kinds of terms enumerated at the beginning of this paper, an analysis of the ordinary meaning of obligability is most apt to yield an inapplicable meaning. An analysis of punishment is most apt to yield an applicable one, for the reason that reflection on non-retributive justification has influenced common-sense beliefs about punishment and hence has influenced conversational language. Since most judgments of the form 'x is responsible' are followed by 'for a crime', and since we usually punish a criminal, it is usually the case that we hold someone responsible in order to punish him. Therefore an analysis of the ordinary meaning of responsibility is apt to yield results similar to punishment. 'Blame' seems to be in a no-man's-land between obligability and responsibility. Ordinary usage has different conditions for the use of these various terms. But however ordinary usage varies from one to another, when we ask "What is the reason for justification for judgments using these terms"? the answer is the same for all: each has an appropriate reformatory effect.

The reformatory effect is different for each. An action which occasions a judgment of obligability also fulfills the conditions for applying the terms 'blame' and 'responsible'. Whenever we say

a person ought not to have done something, we can also say he is to blame for it, or he is responsible for it. However, we do not always say these things. An action, however far from right, deserves to be called obligable. 'Ought not' is sometimes spoken of very lightly as mild reproof, sometimes very strongly, implying that if the words do not have their intended effect, they might well be followed by punishment. The five types of condemnation as first enumerated are in order of increasing force. When a person deserves punishment, all the others apply; obligability is sometimes asserted alone. But however much they differ in force, they are alike in that they are justifiable on the basis of their effectiveness in altering future behaviour.

In order to be effective such judgments need not be directed to the wrong-doer. They may be spoken to others; even so, they are justified because they influence the guilty person indirectly by surrounding him with a "social atmosphere" such that he and others are less likely to commit the condemned crime. This is probably the only reason for clucking at tea, "Old Roger really ought not to have pinched her, you know"!

The principle conditions for the justifiable use of complete condemnation are that the wrong-doer be in such a state that he needs censure and such that censure will have a reformatory effect.[1] The action for which the person is condemned therefore, must have two important causal effects on the person: one negative, one positive. The negative effect is this: the action must not derange or disorganize the personality to such an extent that even the strongest condemnation would have no effect. For example, if a person were to choose to take a potion whose effects made him a raving, incurable murderer, there would be no point in appealing to him or in attempting to rehabilitate him. The positive effect is this: the action must perpetuate a tendency or reinforce a tendency, or start a tendency within the make-up of the personality such that the person is likely to repeat his misdeed in the future. Otherwise there would be no point in condemning a person for a certain wrong action. A person is held responsible, obligable, etc. in a non-retributive sense not because of what he was in the past but because of what he is likely to be in the future. That he is likely to act wrongly is the only reason for our making such judgments against him. We want to affect

[1] Complete condemnations, *i.e.* condemnations which specify a particular action are justifiable mainly by their reformatory effect. They can also be justifiably used to deter others from doing the specified wrong. I shall wittingly ignore the complication of deterrence as irrelevant to the question of freedom.

him so that he will not act wrongly. We hold him responsible after the action ; we condemn him after the action. Hence our only justifiable concern is with the person as he has been made by the result of his wrong choice. Our concern with his former self is justifiable only because without enquiring into what he was like before the condemned action we do not have any way of knowing what he is like after the condemned action. We need to know whether or not he is likely to go astray again, and if he is, how we can prevent it. His former self is dead in the past. Nothing we can do will alter it.

As I have said, the non-retributive sense in which I am now employing the words 'punishable' and 'punishment' are similar to but not identical with that of ordinary moral usage. Only in so far as ordinary usage is non-retributive, preventive, curative are they identical. The sense in which I have been using the phrases 'ought' and 'ought not' are quite different from every-day usage, for the every-day meaning is inapplicable. The only justifiable use of any term of condemnation has to do with the present effects of past actions, not with the actions themselves. We reprove a person in order to reform him. We condemn him for a certain kind of wrong. Also according to ordinary usage we condemn him because he was free to do or not to do the misdeed which either caused or exhibited the tendency to do that sort of misdeed. I shall now have to enquire whether any of the conditions demanded for a 'free act' are reliable signs of the state which warrants condemnation. I have departed from the ordinary usage of terms of condemnation. I shall now enquire how far we need to deviate from the ordinary use of 'free'.

Playing fast and loose with the meanings of ordinary words raises the question of whether, for clarity's sake, one should employ the same word-forms or invent new ones ; and also the question of whether, as long as one has had to modify gravely ordinary usage, he should attempt to preserve any part of it. The first question, as usual, is answered by pointing out that it doesn't make any difference what words we use as long as we make clear what we mean. Also, as is usual with such reflexively repeated dicta, this one is true but never practical to follow. It is hard, clumsy, and confusing to use some words ; easy, elegant, and clean to use others. It would often be best to invent new terms for ethics or meta-morals, *i.e.* for the purpose of discussing moral problems. But for morals itself the terms in use will continue to be used. People do not give up old words easily, or acquire new ones at the suggestion of philosophers. While, ideally, it makes no difference what word-

forms we use, practically we must go on using the old familiar ones. When we have made clear what these old ones should mean in order that they be most usable, we have done our duty in ethics. Ethical thinkers in addition have done their moral duty as ordinary agents when they correct their own usage and try to correct that of others. The second question is simpler. As long as we must in practice use the same old word-forms, then if these words should happen to be partially justifiable (usable), all the more reason for sticking with them, while making it clear that we have dropped some of the conditions defining the term as it is usually used. By a word's being partially justifiable or usable, I mean merely that some of the conditions for its use actually are found *in rerum natura*. And many of the conditions defining the ordinary use of terms of condemnation are fairly reliable signs of the conditions for their justified use.

When the useless conditions for the ordinary use of terms of condemnation are dropped, the corrected, justifiable terms will be used in judgments of the form " You now ought not to do x ", etc. The subject of such judgments is always a person at present after a wrong action has perpetrated its consequences. We can justifiably address such words only to persons who understand and respond to language, and who have a certain minimal amount ofmoral integrity, sensitivity and conscience. I shall assume that this complex set of conditions is to be included in the notion of 'person'. We are now ready to ask whether any of the conditions ordinarily required of a free act are useful signs of the state justifying condemnation. The conditions for the ordinary use of 'free act' are: (1) the act must belong to a person, (2) it must be the kind of action which the person could have chosen or else it must result from desires which could have been subject to choice, (3) somewhere among the causes of the act must be an undetermined choice or desire, (4) the action must not be such as to be caused by the environment and organism of the person, without acting through his personality, (5) it must be determined by a 'pure ego' or something like one, (6) the 'free choice' preceding the act is a datum of experience.

This ordinary notion has been found inapplicable because of (6) about which ordinary usage is confused and (5) which is a pathetic piece of metaphor. If these two conditions are dropped, may the word-form 'free' be justifiably used in the meaning defined by conditions (1) to (4)? In order to answer this question, it is necessary to state exactly what it would mean for judgments of condemnation to be justifiably used. *E.g.*, what useful meaning can be attached to the sentence 'x is obligable because action y

was free'? Its use would be justified under four conditions: (*a*) x as he is now available to us is likely to do actions such as y, (*b*) x as he is now available to us is in such a state that he will favourably respond to condemnation. (*a*) and (*b*) I shall call *factual* conditions. There is also required: (*c*) x as he is now available is known to be likely to do actions such as y, (*d*) x as he is now available is known to be in such a state that he will favourably respond. I shall call (*c*) and (*d*) *epistemic* conditions. The factual conditions exclusively concern the future. To realize the epistemic conditions in our present state of knowledge we must be concerned with the past. We must know about x's past wrongs because we can predict this future only with the help of the laws of psychology which we do in fact know and which tell us that having once sinned he is apt to repeat, and that he has been a morally sensitive person and is apt to remain so. This is the only way we can presently know that when x is condemned he will "see his wrong and have a change of heart". We must know enough about the person to know which kind of judgment of condemnation or what form of punishment will be effective. This depends upon the person. His crime is of interest only in so far as we must in practice consider it in order to know the person.

We do not say, "You ought not to wet your bed" to an average two year old child. We do say, "You ought not to eat spiders". The reason for this is that we have psychological knowledge gained through past observation of children that the first judgment is useless; the second one likely to be effective. We need to know about the past of the individual child in question in order to ascertain whether he is an average two year old. A more perfected state of psychology would enable us to dispense with past observations. In the meanwhile conditions (1), (2) and (4) would seem to be helpful signs of a person's warranting condemnation. They show that the person is likely to respond to moral judgments about his action.

What About Determinism and Indeterminism?

Now, what about condition (3) of the ordinary meaning of 'free act'? This is the requirement that a free act have an undetermined link in its causal ancestry. I think the tool-distinctions needed for answering the question of whether determinism or indeterminism is "required" for condemnation have now been stated with sufficient clarity that the answer easily follows and can be brief.

If we confine ourselves to the factual conditions for justifiable condemnation, the question of whether the condemned action is determined or undetermined is irrelevant. If all our actions were completely caused by our desires, and they by past actions, desires, attitudes, thoughts, etc. which were, in turn, completely caused by our heredity and environment, there would remain exactly the same reason for condemnation that would be present if they were not so caused. Either way will do for moral purposes. However, there is a familiar argument based on the factual conditions for condemnation, which professes to show that determinism is positively required. It states that if a wrong action of a person is uncaused, then the person did not cause it. If the person did not cause it, he is not obligable, responsible, etc. Therefore it would be worse than useless to punish him. Baylis states the argument in clearer, and hence more obviously mistaken fashion. His statement is confined to responsibility, but the context indicates that he would extend the argument to other condemnations.

> When we ask what was responsible for an accident we are asking for the cause of that accident. To say that there is no cause is to say that nothing was responsible. . . . But if our fault was not due to any aspect of our nature, if there were nothing we could have done in the past or can do now which would make us behave better, we are helpless. . . . If indeterminism be true, if some choices are uncaused, then nothing and no one is responsible for them. (p. 63).

Since Baylis holds a completely non-retributive theory of punishment, it is all the more surprising that he identifies what is to be held responsible for a wrong with its cause. Indeterminism is a doctrine about causes, or rather the lack of them, not a doctrine about effects. No indeterminist to my knowledge has stated that a free act has no effects. I think it is tolerably plain that as far as the factual conditions for condemnation go, if (*a*) the effects of an action are dangerous, and (*b*) they could be altered by moral censure or punishment, then condemnation is justified.

The conclusion that determinism-indeterminism is irrelevant follows clearly if we are concerned only with the factual conditions for just punishment. When we add the epistemic conditions a better case can be made for determinism, *i.e.* if we limit the question to the epistemic conditions under the form of our present psychological knowledge. For if we possessed a sensitive brain wave machine which would predict whether a person would commit a crime and what sort of censure or punishment would prevent it, then we would have no need to be concerned with a

person's past. (Of course, we would have to investigate the past lives of other persons, but only in order to accumulate the knowledge necessary to construct the machine.) With the unstated assumption that they are concerned with the justification of punishment in our present state of ignorance, many arguments have been constructed to show that determinism is required. *E.g.*, in the University of California Associates' book, *Knowledge and Society*, the chapter "Freedom of the Will" contains this argument:

> To punish me for the crime of forgery, for example, would be useless unless the punishment tended to deter me from committing acts of forgery in the future. But if voluntary actions are not determined, my future actions are as unpredictable with the punishment as without it. I shall be just as likely to commit the act of forgery in the future no matter if I am now punished for the commission of such an act, or the act is ignored, or I receive a reward for it. But if punishment accomplishes nothing, it is hard to see what one can mean by its justification. Responsibility implies determination rather than indetermination.

The plausibility of this argument depends upon the fact that we must know that punishment will have its intended effect. Therefore, it must have had a determinate effect in the past. It might be added that we must also know that the forger will repeat his crime. In our present state of knowledge, responsibility requires that there be regular sequences of behaviour; one crime leads to another, reform follows punishment. This places a definite limitation on the properties in which certain actions can be undetermined, but the argument seems to jump from the conclusion that reformed acts of choice must be partially determined, to the conclusion that they must be completely determined, and from the conclusion that certain acts of choice must be determined, to the result that all acts of choice must be determined. The conclusion which does follow is this: acts of choice in which the reformed criminal turns away from the action for which he was condemned must be determined in one respect: we must be able to predict that he will not repeat his crime. We need not know which of the many alternatives left to him he will choose.

We must be able to know that condemnation will reform. It must fulfill its intention to prevent repetitions of the wrong action for which the person is reproved. If we are to know this it must be because we have observed that past condemnations have reformed. Therefore, we must possess on the basis of these observations many psychological laws which are specializations

of the following form: "Persons of character P who have been affected by a wrong action of type Q will not repeat Q after condemnation *p*". Complete determinism is hardly required in order that we might learn such laws, nor even determination of acts of choice which lead to wrong actions where the person has not been correctly condemned for that action. Arguments for determinism are based on the presumption that we must be able to predict all of the consequences of a condemnation. We need to know only that it will prevent the wrong action in question—and, it should be added, not cause wrong actions of another kind.

Now, weak as the determinist's argument is, if it is based strictly on our present state of knowledge, it is weaker still. For we do not know the laws which are required to justify punishment except in a very general and incomplete way. In the strong sense of 'justification' which I have been using, punishment is never justified. We are forced to take action on the best knowledge available; attempted reform, more often than not, does not come off. And we never know whether our failures are the result of cussed indeterminism or are due to a lack of knowledge of the deterministic laws. But this is small comfort to the indeterminist, for on the possibility of finding these laws, he, too, must insist.

I want to conclude with a reminder that my topic has been a rather limited portion of the "free-will question". I have been concerned only with the moral requirements for determinism or indeterminism. I believe it may be concluded that the practice of moral praise and blame places insignificant limitations on the kind of indeterminism which is defensible, but does not require either determinism or indeterminism. In view of this, the more important arguments for determinism or for indeterminism are of another sort, and in all historical fairness it must be stated that they have recently been of another sort. Arguments for determinism have been taken from an analysis of scientific method, arguments for indeterminism from an analysis of time (the so-called "process" view of Bergson and Whitehead).

FRANK B. EBERSOLE.

Oberlin College, Ohio.

VIII.—LUDWIG WITTGENSTEIN, 1934-1937

By John Wisdom

I have no notes worth speaking of about what Wittgenstein was saying in the years 1934, 1935, 1936 and 1937, when I attended his lectures and he talked to me about philosophical questions. I am relying on memory. As I remember, when I first went to a lecture by him, he was talking about the question, "What is it to understand a general term, such as 'plant'?" (See the first pages of the Blue Book.) He studied the cases in which we say of someone, "He understands", "He knows the meaning", "He is being taught the meaning", "After all he doesn't know the meaning", "I know the meaning", and so on. This led to his emphasizing the point he expressed by, "We have the idea that the several instances of a general concept all have something in common" (See the Blue Book.) (The concepts of 'understanding' and of 'having a meaning' are no exception.) He said that in applying the same word to several instances we mark a family resemblance—not the possession of something in common (as all ticket holders possess something in common—a ticket which matches a ticket I may hold in my hand as a sample, or as all alcoholic drinks possess something in common, in that from all of them may be distilled the essence of an alcoholic drink—alcohol).

This remark of his—that in applying the same word to several instances we mark a family resemblance and not the possession of something in common—was connected with a point which on one occasion at the Cambridge Moral Sciences Club he expressed in the words, "We have the idea that the meaning of a word is an object". This is connected with his saying "Don't ask for the meaning, ask for the use", recommended at the Moral Sciences Club as a supplement to "The meaning of a statement is the method of its verification". And all this is connected with the question "What happened when you understood?", and thus with his study of "What happened while I was expecting so-and-so from 4 to 4.30?", and so with how much a question as to what happened when someone understood, believed, remembered, was reading, was coming to a decision, felt frightened, etc., is a matter of what happened before, and what will or would happen after, he understood, believed, etc., and with how our recognition of this is hindered by 'the idea of a mental mechanism', the hidden movements in which are these activities of the mind.

"We have the idea that the meaning of a word is an object" is also connected with "The application (every application) of every word is arbitrary". And this is connected with the question, "Can you play chess without the queen?". (If I were asked to answer, in one sentence, the question "What was Wittgenstein's biggest contribution to philosophy?", I should answer "His asking of the question 'Can one play chess without the queen?'".) And all this about understanding is connected with his study of what it is to prove a thing, with the fact that people were often exasperated by his ending the discussion of a philosophical puzzle with "Say what you like", with his saying to me on one occasion when I spoke of an unsuccessful philosophical discussion "Perhaps you made the mistake of denying what he said", with his saying "I hold no opinions in philosophy", and with his saying that he didn't *solve* philosophical problems, but *dissolved* them. At the same time he was always anxious to make people feel the puzzle—he was dissatisfied if he felt they had not done this.

The idea that the several instances of a general concept all have something in common is connected with the craving for a definition, "the idea of an exact calculus", the model, the analogy, of an exact calculus. The fascination of this model for our language is connected with the fascination of models suggested in our language—the idea that the soul is a little man within, the model for our minds of the closed picture gallery, the model for causation of the wire connexion.

The substitution of the family resemblance model for the property-in-common model, and the substitution of "Ask for the use" for "Ask for the meaning" is linked with the procedure of explaining meaning by presenting not a definition but cases, and not one case but cases and cases. And this is linked with dealing with the philosophical, metaphysical, *can't* by presenting cases and cases.

Thus Wittgenstein said that if someone says "One can't know the mind of another", one may ask "Would you call this knowing the mind of another, or that, or that?" For example, one may ask "Suppose a nerve of your body was joined to a nerve of Smith's, so that when someone stuck a pin into Smith you felt pain, would that be knowing, having, Smith's pain?" By this procedure, *either* (1) it is made clear what we would call 'knowledge of the mind of another', and hence what it would be to know the mind of another, *or* (2) it appears that we can't imagine what it would be to know the mind of another, and that not merely in the sense in which few people can form an image of a

creature with twenty-two legs on one side and twenty-one on the other, but in the sense in which there is nothing the speaker would call 'knowing the mind of another' or 'feeling the pain of another'. If (1) is the case, then the person who says "One can't know the mind of another" is saying of something which conceivably could happen, that it doesn't. Pigs don't fly. But if (2) is the case, then he has taken away the use of 'know what is in the mind of another' which we have been taught, and not provided us with a new use. And this makes meaningless his question "Are other people automata?".

I do not mean that such a procedure constituted Wittgenstein's whole treatment of the puzzle about knowledge of the mind of another. By no means. On the contrary, he emphasized the fact that in teaching the child the use of 'in pain' we not only point to others who are moaning, perhaps, but also pinch the child until it hurts, and say "That's pain". And with this two-fold teaching and learning goes a two-fold method of verification which makes "He has toothache" very different from "It goes in jerks".

If one says that Wittgenstein showed metaphysical questions to be meaningless one must remember that he also said that what the Solipsist *means* is right. One must remember if one asks "Can one attach a meaning to 'Can one ever know that Smith is in pain, that there isn't a white rabbit between any two articles of furniture in this room'?" that one also asks "Can one play chess without the queen?".

Cambridge University.

IX.—DISCUSSIONS

Mr. GEACH ON RIGOUR IN SEMANTICS

In a recent note,[1] Mr. P. T. Geach offers a series of criticisms of Carnap's treatment of semantics. He aims to prove that this discipline can be no good since one of its leading exponents makes so many blunders. His conclusion is that it may perhaps after all not be so foolish to undertake 'direct analysis' in language everybody can read.

I must confess that Mr. Geach's argument appeals to me, having myself, some time ago,[2] used a similar one and tried to show that the direct method of language analysis as practised by many contemporary British philosophers is not reliable since one of its leading exponents made on one single occasion so many blunders.

Though the arguments are similar, indeed, it seems to me that it would not be difficult to point out the essential differences between a mistake within a rigorous discipline and a mistake in interpretation of "correct" language. But, fortunately enough, no such effort on my side is necessary, as Mr. Geach's argument is almost completely unjustified. My discussion will follow Mr. Geach's criticisms point by point. Carnap's *Logical Syntax of Language*, *Introduction to Semantics*, and *Meaning and Necessity* will be referred to as LS, IS, and MN, respectively, Mr. Geach's note as RS.

(1) Mr. Geach accuses Carnap of using dots and dashes as a device to avoid certain difficulties. Carnap himself explains this use quite clearly: "The context is indicated only by dots instead of by a second-level variable, in order to make the definition applicable also to systems not containing such variables." (MN, p. 147, n. 5.) Mr. Geach argues: "But these dots *are* a sort of variable; only it is left obscure whether they are, after all, a second-level variable of the object-language, or are rather a variable of the meta-language".

The facts are simple and obvious. The dots and dashes *function*, in this case, as second-level variables of the object-language (and *function*, in other cases, as sentence-variables or as many other things) but they *are not* second-level variables. Their function is explicitly given or implicitly understood in the meta-language. It may be that certain difficulties are connected with such usage. But Mr. Geach has not succeeded in pointing them out, nor has he succeeded in showing another equally simple, if not simpler, way of transcribing Russell's definition of the class-symbol into a language-system which does not contain second-level variables, and such

[1] "On Rigour in Semantics", MIND, October 1949, pp. 518–522.

[2] In "Analysis of 'Correct Language'", MIND, N.S., vol. lv., 1946, pp. 328–340.

systems are frequently studied. (By the way, there are two inaccuracies—or misprints—in Mr. Geach's rendering of Carnap's 33.2.)

(2) Mr. Geach accuses Carnap of giving two inconsistent definitions of the sign of definition on pages 17 and 20 of IS, namely

(2 : 1) A definition has the form '. . . = D*f* – – –'; this means: "'. . .' is to be an abbreviation for '– – –'",

and

(2 : 2) '= D*f*' is to mean 'is (hereby defined to be) the same as' or 'if and only if'.

It is hard to understand why Mr. Geach calls these explanations belonging to a brief survey given of certain elementary parts of symbolic logic 'definitions'. At the end of (2 : 1), Carnap refers to his § 24, where 'definition' is more extensively dealt with, though still not defined, in any rigorous sense of this word.

But what is "inconsistent" about these two explanations? I quote:

Now 2 : 2 expounds the sign '= D*f*' in terms of an expression, 'if and only if', which can stand only between sentences or sentential functions; 2 : 1 expounds it in terms of the expression 'is to be an abbreviation for', which can stand only between *names* of expressions, not between sentences or sentential functions.

So what? At the best, we have here one more instance of Mr. Geach's failure to observe the distinction between the *use* and *mention* of expressions.[1] The reader will verify that 'is to be an abbreviation' appears in Carnap's first statement (without quotes), but "'if and only if'" in his second statement (with quotes), so that Mr. Geach's putting them on the same foot by using the same expression 'expounds it in terms of the expression' is more than misleading.

(3) Mr. Geach's third accusation boils down to this: Carnap regards the expression 'Des_G ('drei', three)' as a sentence of the meta-language he uses in his § 12 of IS; "but this is certainly not an English sentence".

Well, we need only reproduce Carnap's words in the first passage of § 12:

We use as meta-language in this section the English language supplemented by variables. . . . Instead of 'u designates v in S' we write 'Des_S(u,v)'. . . .

'Des_G('drei', three)' *is* a sentence of (supplemented) English, since introduced as an abbreviation of a sentence in what is doubtless good English. 'Des_G('drei', three)' is hence an example of an English sentence having the form $\mathfrak{pr}_i$ ($\mathfrak{A}_j$, $\mathfrak{A}_k$).

This is the kind of argument by which Mr. Geach tries to show that "there is a recurrent error in Carnap's use of letters"!

[1] Cf. the note of R. M. Martin, "Mr. Geach on Mention and Use", MIND, October 1949, pp. 523–524.

(4) Mr. Geach is indeed right in pointing out that the definition given by Carnap for the adequacy of a predicate $\mathfrak{pr}_i$ in M for the concept of truth with respect to an object language S holds good only in the case that M is English (supplemented by certain variables). But this point is so trivial that no reader (with a certain minimum of understanding and good will—and such a minimum is required even by the most formal and rigorous expositions !) should feel any unhappiness about it ; he will surely be able to make the necessary alteration for the case where M is some other language or does not contain variables.

(5) Once again Mr. Geach is right. Carnap made indeed two grave mistakes.

(*a*) His use of quotes around syntactical variables or variable expressions *might* mislead. But I am sure nobody will be misled. Carnap himself has already pointed out (IS, 7) an exact method of dealing with cases where misunderstandings are possible.

(*b*) There is really a misprint on p. 22 of LS. A pair of quotes are missing. They are missing in the English translation only ; in the German original, they appear at their proper place. Now, there are very few misprints in Carnap's publications, and we are grateful to Mr. Geach for having detected one more. But we have already shown our gratitude by pointing out *two* misprints in his article. There are, by the way, a few more, and to one of them we shall call attention later.

(6) Mr. Geach simply has not grasped the full import of Carnap's distinction between variables, constants of undetermined meaning, and constants of determined meaning. If ' B ' is an abbreviating constant of an object-language with undetermined meaning, then it *makes sense* to write " for any ' B ' " in the meta-language, and the sense is very obvious.

(7) Mr. Geach criticizes Carnap for introducing the null thing a_0 as the common descriptum for those descriptions which do not satisfy the uniqueness condition (a_0 is the descriptum, by the way, and not ' a_0 ', as Mr. Geach has it on p. 521 of RS, line 4 from below, a rather curious " misprint ").

He compares this convention with the following " really similar mathematical " one :

It is possible to count among natural numbers also the *standard number*, characterized as that finite integer which is the least common multiple of all natural numbers. . . .

But is Mr. Geach indeed unable to see what must be a very obvious difference ? It can easily be proved that no natural number is a common multiple of all natural numbers ; but can Mr. Geach, or anybody else, prove that there is no thing which is part of every thing ? Perhaps he can, starting from certain definitions of ' thing ' and ' part of thing ' which are suitable to him. But this does not prove that it is impossible to show the existence of the null thing

from other definitions which are not less suitable, to other persons and for other purposes.

Mr. Geach's comment, "in plain English", on Carnap's making the null thing correspond to the null class of space-time points as describing it "as existing nowhen and nowhere" speaks for itself.

There are, of course, serious reasons for *not* introducing a null thing, but these are of a philosophical and not logico-semantical nature.

Y. BAR-HILLEL.

Jerusalem.

THE PREDICTION PARADOX

In Mind, 1948, Mr. O'Connor published an article entitled "Pragmatic Paradoxes". He there gave the following case: "The military commander of a certain camp announces on a Saturday evening that during the following week there will be a 'Class A blackout' . . . an exercise which the participants cannot know is going to take place . . . it follows from the announcement of this definition that the exercise cannot take place at all. It cannot take place on Saturday because if it has not occurred on the first six days of the week it must occur on the last. And the fact that the participants can know this violates the condition which defines it. Similarly . . . it cannot take place on Friday either, because when Saturday is eliminated, Friday is the last available day". If Friday is eliminated, Thursday will be eliminated for the same reason, and so on back through the week, with the result that the blackout cannot occur at all.

I hope Mr. O'Connor will not mind my giving his paradox the new and somewhat more appropriate name of "the prediction paradox". I hope, too, he will not mind my using a different illustration of a more practical, poignant and permanent interest than his, and that he will excuse my refining his formulation a little.* (Mr. O'Connor so states the paradox that it is possible for the announcement to be rescinded, with the consequence that the non-occurrence of the blackout rather than its occurrence could be predicted.)

The Paradox: A headmaster says, "it is an unbreakable rule in this school that there be an examination on an unexpected day." The students argue that the examination cannot be given on the last day of the school year, for if it had not been given until then, it could be given only on that day and would then no longer be unexpected. Nor, say they, can it be given on the next to the last day, for with the last day eliminated, the next to the last day will be the last, so that the previous argument holds, and so on and so on. Either the headmaster gives the examination on an expected day or he does not give it at all. In either case he will break an unbreakable rule; in either case he must fail to give an examination on an unexpected day.

The paradox arises I think due to a confusion between two meanings of the term "or", a collective and a distributive. When we predict we refer to a range of possibilities which are as yet undistinguished one from the other. They are connected by means of a collective "or", prohibiting the separation of any one of them from the others, without the introduction of some power or factor not included in the concept of the range. Since predictions always refer to a range and never to the specific determinations of it produced in fact, the predictions must be supplemented by history or the imagination if we are to

*I owe the illustration to George B. Burch.

select and eliminate first one and then another alternative. What is selected and eliminated in history or in the imagination will be something distinct, focussed on, actualized, connected with others by means of a distributive " or ". If we avoid confusing these two meanings of " or ", our paradox, I think, will disappear. Perhaps this will be somewhat clearer from an examination of the following five cases.

1. On the last day of the term, the headmaster says, " Oh, by the way, I forgot to tell you about an unbreakable rule in this school. We must have an examination on a previously unexpected day. Please take out your pencils and write down and answer the following questions." There is no doubt but that the day and the examination will be totally unexpected. Instead of prohibiting the giving of the examination, as Mr. O'Connor seems to infer it must, the announcement will require the examination, to the consternation of totally unprepared boys. If we wish to produce a paradox we must suppose that the announcement be made not on the last day but on some other. This points up the fact that the paradox is a function, not of the announcement, but of the relation of the announcement to the announced occurrence.

2. The headmaster, on some day of the term other than the last, announces : " There is an unbreakable rule that we must have an examination on a previously unexpected day, and this is it ". Once again the day will be unexpected; once again there will be no paradox. If we wish to produce a paradox we must insist that the announcement be made before the day on which the examination is given. The first case discussed above is evidently a special case of this; the general rule is not that the announcement must be made on a day before the last, but that it must be made on a day before the examination is given.

It might be argued that if we were to treat the announcement as definitory of a range of occurrences, determining what events we are thereafter to consider, the unexpectedness will disappear. On this view the element of unexpectedness relates solely to the act of announcing a rule, and has nothing to do with what in fact takes place under the aegis of that rule. If we could eliminate the element of unexpectedness involved in the act of announcing we should then have nothing unexpectable left. The position cannot, I think, be maintained, for firstly, we cannot eliminate the element of unexpectedness in the act of announcing the rule, and if we could eliminate it, there will still be left an unexpectedness connected with the occurrence of the examination.

Instead of meeting the question at issue, the above gives it a new location. It leaves us with the question : can we expect the announcing of an announcement which says that it will be announced unexpectedly ? To this no answer is given. Yet it seems evident that we have no grounds in an announcement for an expectation that the announcement will be announced in the particular way or time

in which it is announced. There is something unexpectable in the act of announcing, just as surely as there is something unexpectable in the occurrence of the events covered by the announcement. Between a possible and an actual announcement there is a distance. He who starts with the first will find that the second comes unexpectedly, and for the same reason that the fulfilment of an announced rule comes unexpectedly—because it is a particular not knowable from the vantage point of what, with respect to it, is only a possibility.

If we have only one day left on which an examination can be given we know the day of the examination, but we cannot know or predict its length, its tediousness, etc. These are subordinate specifications of the examination and depend for their existence on the giving of the examination. I can say that I will positively do 'something or other' tomorrow and yet not know positively what it will be. I can be positive that "this or that will be", *i.e.* of a range of possible occurrences connected by a collective "or", which allows no one of those occurrences to stand apart or be distinct from the others. This does not make it possible for me to know that just this will be done, or that something other than this will be done. There is a great difference between "it is true that either x or non-x is the case", and "it is true that x is the case or it is true that non-x is the case",[1] between $f(x \vee \bar{x})$ and $f(x) \vee f(\bar{x})$. The former is a necessary truth, the latter is true only when one has isolated the x and the non-x, an act which requires one to leave the realm of possibility for the realm of time, history, becoming. To become is to convert the collective into the distributive "or".

In the above two cases we suppose that there is an actual announcement of a rule, and that there is only one case to which the rule applies. The actual announcement comes unexpectedly and so does the application of the announcement. Let us ignore the first unexpectedness and concentrate on the second. The transition from a rule announced to an instance of it requires us to go beyond the conditions expressed in the rule. It makes no difference if the rule have many cases or only one case to which it can apply. To put it still another way, $f(x)$ is not deducible from $f(x \text{ or } y)$ even when it is the case that y is 0.

In the above two cases what is unexpectable—or unpredictable or undeducible, if one likes—is both the occurrence of the only available instance of a rule and the announcing of the rule. In the following cases we take the rule for granted, thereby removing what was a source of confusion in the previous cases. The question that then remains is whether or not one can deduce that there cannot be an unexpected examination inside the time-span which the announcement defines and limits in advance. It is my contention that there can be such an unexpected examination. This is a consequence of the fact that the announcement makes use of the collective "or", and that the examination occurs only when detached from other events and times by means of the distributive "or". The occurrence

[1] *Cf.* Aristotle's, *De Interpretatione*, 18b ff.

of the distributive "or" and the items it disjoins are always unexpectable from the standpoint of what is constituted by the collective "or".

3. The headmaster makes his statement in the evening, two days before the last. It is then true that "the examination will be given the next day or the day after"; it will not then be true that "it will be given the next day" or true that "it will be given on the last day". There is as yet no distinct next day or day after on which the examination could be given. If the examination is not given the next day, when that day comes around, there will, of course, be no alternative left but that of the last day. It is no surprise then that if we think of ourselves as at the last day, as the paradox requires, and then proceed to eliminate that day, we will find there is no day left on which the examination could be given, the other days having already been eliminated in the act of coming to the last day. When the announcement is made it is not predictable which of two days will be the day on which the examination will be given, but once we have eliminated the first of these days in fact or in imagination, the examination must occur on the second day. But until the days have been distinguished they are not alternatives, excluding one another. He who supposes that they are, and therefore argues that the examination cannot be given, will find to his surprise that the examination is given and at a totally unexpected time.

4. Suppose the examination had not been given by the time the last day came around. Suppose, too, that a student had paid no attention to the announcement when it was made some time previous to the last day. Suppose, finally, that the student suddenly remembered the announcement that day. The examination will certainly come on an unexpected day for him, though there is, in fact, only one day on which it could occur. This conclusion does not go counter to our observation above. A student to whom the announcement was not communicated (because of his inattention or otherwise) is confronted with the possibility not of "examination today or not today" but with the possibility "examination or no examination today". He does not know, before he remembers the announcement, whether an examination will be given or whether an examination will not be given. In remembering the announcement he eliminates, as no longer pertinent to him, the alternative "no examination today". The elimination is unpredictable; it cannot be determined by a consideration of the range of possibility open to him. When, though, the alternative of "no examination today" has been eliminated by him, there is, of course, no alternative left but that of "examination today". What will then be unpredictable will be the exact time or nature of the examination. This he can know when he is in the actual presence of the examination, not before.

When and as the examination occurs it is not unexpected; it is unexpected only before it occurs. Or, if by "examination" we mean to represent a possibility of long or short, fair or unfair, etc.,

examination, an examination of some sort will be rightly expected but one could not predict its length or fairness—prediction being understood to be grounded in present data and to infer, with warrant, what this entails.

5. If the headmaster were to make his statement before school opened, the students would have to be prepared for an examination on any day at all. That examination will occur some time in school term. But they cannot know, they cannot predict, they cannot rightly expect the day on which it will occur. If the examination is not given by the time the last day comes around, the examination will have to be given on that last day. That day will have been unexpected all along. When it comes around it will, of course, no longer be an unexpected day, for what is now present is no longer unexpected. But the last day cannot be dislocated from the other days at the time the announcement is made, expect by a kind of theoretical anticipation of actual history, and thus by moving away from the announcement of a possibility to that state of affairs where possibilities are specified and distinguished one after the other. The announcement tells us that some one day will be selected as the day of the examination. Whatever one it be, it will be unexpected, so far as we view it from the standpoint of the announcement; it will be expectable so far as we have eliminated the others and thus made it distinct from them—a status it does not have in the announcement.

In all the above cases the students can be certain that some one of a number of alternatives in a given range must occur. They can know with surety the range as a tissue of possibilities connected by collective "or"; they will not then know the items in that range so far as these are connected by a distributive "or". The range might be a range of dates for a possible examination, in which case the students can be certain that an examination will occur but cannot be certain as to just when it will occur. Or it might be a range of events which can be given the common name "examination or no examination". In such a case it will be certain that something will occur, but it will not be certain as to just what this will be. Should we succeed in eliminating all alternatives but one within a given range, there will be no uncertainty regarding the necessity for the remaining alternative. The elimination of that alternative will leave us with no alternative at all; it will involve the cancellation of the entire range.

The paradox of prediction arises because it is forgotten that there is no distributive "or" until we do something to the range of items connected by a collective "or", or more concretely, because we forget that the dates on which an examination can occur are, on the date of an announcement made before the last day, not yet distinguished, and that we have already eliminated all other days when we treat the last day as a day distinct from all the rest.

PAUL WEISS.

Yale University.

A NOTE ON REICHENBACH'S CLASS CALCULUS[1]

In his recent work, *Elements of Symbolic Logic*, Reichenbach has devoted a chapter to the Calculus of Classes. Choosing a few points from this chapter, I shall attempt to elaborate a criticism applicable to the book in general. The point of the criticism is merely this: Reichenbach has mixed his own philosophical commitments with a formal explication of a logical system; a system that (to me) should be neutral with respect to these extra-logical issues. Thus, it is with an eye to the separation of non-formal tenets from issues primarily formal that we turn to Reichenbach. But it is not only this we shall pursue, for, to make the criticism pointed, I shall take representative commitments of his and contrast them with differing views that seem equally compatible with the formalism he presents.

Proceeding in customary fashion, Reichenbach first defines the class symbol and the "relation" of class membership, by the use of a propositional function. Stepwise, the calculational constants 'disjunction', 'conjunction', 'complementation' and 'identity' are introduced by definition. By such rewriting rules, adding the functional operators and connectives, one obtains a *PM*-type system (restricted to one term predicates of the first level) written in the notation "proper" to classes. With this brief introduction, I leave the programmatic feature to consider Reichenbach's remarks about them. Specifically, I wish to comment on (1) the treatment of the extensional definition of classes, (2) the remarks on the identity of classes, (3) the remarks on the method of "extensive abstraction".

When one defines a class, as above, by a propositional function he has, in Reichenbach's terms and in general usage, been said to define the class intensionally. However, it is clear that one might define a class, if it be finite, by enumeration of the members said to comprise the class. Such a definition, whose right-hand member contains no descriptive predicates but only particulars, Reichenbach calls extensional. The extensional definition can, Reichenbach claims, always be replaced by an intensional definition. If this be so it is a formal feature and not our present problem. Rather, what is our problem and of present interest are the remarks Reichenbach makes concerning this feature. We can, he tells us, regard the fact that the name of an object is placed on a list (enumerating the members of a class) as a property of that object—that is, as the defining property (p. 193). Reflection, however, shows that this need not be admitted; that it might be claimed that the defining property is not, nor ever can be, that the object is named, say, "on this sheet of paper". If one distinguishes a name, what it names, and the thing named, then clearly this confuses the thing named and the name of the thing. For, I should claim that the object is not "on this sheet of paper",

[1] Reichenbach, H., *Elements of Symbolic Logic*, Macmillan, New York, 1947. All page references are to this work.

so "being on this sheet of paper" cannot be a property of it. Yet the name of the object is not the object itself, so its (the name's) being mentioned in a list cannot be a defining property of its referent either.

In the introduction of the identity of classes we find ourselves confronted with a somewhat unusual notation (p. 199). Two classes, we are told, may be said to be identical if they have the same members. We must, however, be careful. All that is here required Reichenbach continues, is that the extensions coincide; this does not presuppose that the classes are determined by functions having the same meaning. In the former case, where the equivalence between the members of the two classes is "not a tautology", he suggests we may speak of extensional or "adjunctive" identity. In this case, all members of the one class *are* members of the other, as for example, all human beings are as a matter of fact featherless bipeds. In the latter case, however, the identity is "connective". For example, the class of even prime numbers is connectively identical with the class '2', and all members of the one class *must* belong to the other. It is not clear whether Reichenbach here seeks, at least implicitly, a possible resolution of the so-called paradox of analysis, the distinction here offered in class terms. Whether or not this is the case, there is not, in general logical parlance, an adjunctive identity and connective identity. Rather, in the treatment of classes there is only one identity; that is, two classes are said to be identical if the functions that determine them are coextensive, whether or not the statement of "this equivalence is a tautology". Whether or not the identity is adjunctive or connective is a second and different question.

It is in section 37, "the principle of abstraction", that Reichenbach appears most clearly to lapse from a formal presentation of the class calculus. For it is in this section that Reichenbach, impressed by the "principle of abstraction", proffers remarks I can only regard as a treatment of the historic realism-nominalism issue. For these remarks, which we shall shortly examine, it seems clear that Reichenbach finds in the method of extensive abstraction a resolution of the hoary problem; a resolution in favour of nominalism. The method of extensive abstraction, first introduced by Russell and Whitehead, reflects the possibility of construing or introducing certain terms by definition. Although this method has usually been stated in terms of defining classes from properties it need not be so restricted. One could well define descriptive terms (as properties *or* as classes) from basic terms, and similarly one may introduce by definition logical concepts in terms of basic logical symbols. Likewise it would be possible to define a descriptive class name in terms of basic descriptive class names. Accepting this notion of extensive abstraction, it is evident I think that it has nothing to do with the question of the existence of universals; or, to put it linguistically, with the need of basic descriptive terms other than constant particulars.

Having attempted to explicate briefly the irrelevance of extensive abstraction to the nominalism issue, I must now try to make good my claim that Reichenbach does think it relevant. But consider the following : we need not, Reichenbach tells us, regard the property as something removable from the concrete object (p. 209). Rather the notion of a property can be replaced by the notion of class. For, in the case of a rose a given shade of red, we can define the class of objects colour-similar to the given rose, and speak of an object as a member of this class instead of saying it is red. Such a procedure, he continues, shows that *abstract* properties are reducible to *concrete* relations, and has the merit of eliminating unnecessary entities in Occam's sense (p. 210). These remarks give an accurate indication as to the tenor of Reichenbach's claims, and I hope, make good my claim. ' Red ' may, of course, be defined as the property of being colour-similar with " this rose ", or as the class of all objects colour-similar with " this rose ". But, apart from what may be meant by speaking of the relation as " concrete ", and the definition as "ostensive " (p. 210), there remains the fact that a realist might well claim that this procedure does not get rid of the universal ' colour-similar '. What it all seems to come down to is the possibility, in principle, of defining colour words in terms of a proper name and the basic descriptive relation of colour-similarity.

If one accepts that formalisms are neutral with respect to philosophic issues, then it seems clear that Reichenbach has mixed two levels of discourse. It need not, of course, be a confusion to do so. Yet, to do so consciously but not explicitly makes difficult an already arduous subject matter.

ROMANE CLARK.

State University of Iowa.

DESCRIPTION AND THE ANTINOMY OF THE NAME-RELATION

THE following is an example of what Rudolf Carnap has called " the antinomy of the name-relation ".[1] Consider the following sentences:

(1) Scott = the author of *Waverley*.
(2) George IV wished to know whether Scott was the author of *Waverley*.
(3) George IV wished to know whether Scott was Scott.

Sentences (1) and (2) are presumably true statements of fact, and sentence (3) is presumably not a true statement of fact. Yet sentence (3) seems to follow logically from (1) and (2) in accordance with a principle which Bertrand Russell has formulated as follows:

(4) If a is identical with b, whatever is true of the one is true of the other, and either may be substituted for the other in any proposition without altering the truth or falsehood of that proposition.[2]

This is the same principle which W. V. Quine has called the principle of the *indiscernibility of identicals*. He has formulated it (more accurately) as follows:

(5) Given a true statement of identity, one of its two terms may be substituted for the other in any true statement and the result will be true.[3]

The above puzzle is listed by Russell as one of three which any adequate theory of denoting ought to be able to solve.[4] The solution proposed by Russell is well known. He says that the inference to (3) from (1) and (2) is not warranted by the principle (4) because when we eliminate the descriptive phrase from (2) in accordance with his definition we find that the resulting sentence contains no constituent of the form " $(\imath x)(\phi x)$ " for which we may substitute " Scott ".

In fact, the descriptive phrase in (2) may be eliminated in two ways according as we interpret it as having a *primary* or a *secondary* occurrence in that sentence. In the first case we interpret (2) as being of the form:

$$(6) \qquad [(\imath x)(\phi x)](\psi(\imath x)(\phi x))$$

and the result of the elimination is:

$$(7) \qquad (\exists c)[(x)(\phi x \equiv (x = c)) \,\&\, \psi c]$$

[1] Rudolf Carnap, *Meaning and Necessity*, section 31.
[2] Bertrand Russell, " On Denoting ", reprinted in *Readings in Philosophical Analysis*, ed. H. Feigl and W. Sellars, p. 108.
[3] W. V. Quine, " Notes on Existence and Necessity ", *The Journal of Philosophy*, vol. 40, p. 113.
[4] *Op. cit.*, p. 108.

In the second case we interpret (2) as being of the form:

(8) $\chi\{[(\imath x)(\phi x)](\psi(\imath x)(\phi x))\}$

and the result of the elimination is:

(9) $\chi\{(\exists c)[(x)(\phi x) \equiv (x = c)) \,\&\, \psi c]\}$

An English version of (7) is:

(10) One and only one individual wrote *Waverley* and George IV wished to know whether that individual was Scott.

An English version of (9) is:

(11) George IV wished to know whether one and only one individual wrote *Waverley* and that individual was Scott.

Now neither (10) nor (11) contain the phrase "the author of *Waverley*" and thus Russell's solution will apply no matter whether we interpret that phrase as having a primary or a secondary occurrence in (2).

Unfortunately, however, there seem to be good arguments against the acceptance of either of the interpretations which Russell suggests as possible. First let us consider the interpretation which renders (2) as (7), *i.e.* the interpretation which accords "the author of *Waverley*" a primary occurrence in (2). In accordance with this interpretation (3) is a logical consequence of (1) and (2) by substitution into the theorem of *Principia Mathematica:*

$$14.15\ \{(\imath x)(\phi x) = b\} \supset \{\psi(\imath x)(\phi x) \equiv \psi b\}.$$

This result alone, it seems to me, would be sufficient grounds for rejecting this interpretation of (2). But Quine has given a general argument which makes it seem plausible that sentences of the kind (10) should not be accepted as meaningful at all.[1] For, (to paraphrase his argument) who would that individual be of whom George IV wished to know whether Scott was he? Would that individual be the author of *Waverley*? But such an affirmation will be at once true in the form (2) and false in the form (3).

Finally, a third objection against the interpretation of (2) as according a primary occurrence to "the author of *Waverley*" is that under this interpretation (2) has the consequence:

(12) One and only one individual wrote *Waverley.*

It seems to me that (2) could have been true even though (12) were false. A sharper form of this objection is obtained through a consideration of the following:

(13) Harry Truman wished to know whether Bertrand Russell was the author of *Principia Mathematica.*

According to the interpretation being considered this implies:

(14) One and only one individual wrote *Principia Mathematica.*

Again, it seems to me, that (13) might well be true even though (14) is, of course, false.

[1] *Op. cit.*, p. 124.

Let us turn now to the alternative interpretation which accords a secondary occurrence to "the author of *Waverley*" in (2). According to this interpretation (2) would have the same meaning as (11) (or at least be logically equivalent to (11)). But in order to see that (2) could be true and (11) false, we have only to consider that George IV was (no doubt) unacquainted with Russell's theory of descriptions. In order to construct a sharper form of this objection, consider the following sentences:

(15) George IV believed that Scott was the author of *Waverley*.

(16) George IV believed that one and only one individual wrote *Waverley* and that individual was Scott.

According to the interpretation which we are considering, (15) and (16) have the same meaning (or are logically equivalent). But it seems to me entirely possible that George IV should have entertained the one belief without entertaining the other. If we accept the analysis of belief-sentences which has been offered by Carnap,[6]

[6] *Op. cit.*, p. 62.

sentence (15) would be interpreted by:

(17) There is a sentence S_1 in a semantical system S such that (*a*) S_1 is intentionally isomorphic to "Scott was the author of *Waverley*" and (*b*) George IV is disposed to an affirmative response to S_1.

Sentence (16) on the other hand would be interpreted by:

(18) There is a sentence S_1 in a semantical system S such that (*a*) S_1 is intentionally isomorphic to "One and only one individual wrote *Waverley* and that individual was Scott" and (*b*) George IV is disposed to an affirmative response to S_1.

On this (plausible) view concerning the interpretation of belief-sentences, it seems perfectly clear that (15) and (16) are not logically equivalent.

LEONARD LINSKY.

University of Illinois.

X.—NEW BOOKS

Language and Intelligence. By JOHN HOLLOWAY. London: Macmillan and Co., Ltd., 1951. Pp. xv + 189. 12s. 6d.

THIS book is an interesting attempt to map out the relations between the understanding use of signs and intelligent behaviour generally. Mr. Holloway thinks that both have been misinterpreted by being too closely connected with each other, and also by being too closely tied up with peculiar inner *experiences* of interpretation and understanding, which don't invariably accompany them, and which aren't at all reliable criteria of their presence. He holds that it is in *performance*, whether actual or potential, that intelligence is to be sought for and pinned down, and that speech is but a special sort of human performance, which isn't invariably intelligent, and which is in other respects quite as varied as the non-linguistic performances out of which it develops. Parts of Mr. Holloway's book run on lines familiar in psychological controversy, while others summarize views held by modern linguistic philosophers. But some of his arguments are both novel and ingenious, and he shows a sound disposition not to deny, ignore or distort those data of the inner life whose importance, for the particular analysis he is undertaking, seem to him somewhat questionable.

The first two chapters of Mr. Holloway's book contain his criticisms of the traditional mentalistic accounts of meaning and understanding: to me they seem the most interesting part of his work, though I also find them the most open to criticism. He deals with the views of the British empiricists, to whom symbols acquire meaning through their association with images, and then goes on to deal with views which make meaning accrue to symbols through their association with 'concepts' and universals. He passes rapidly over Locke's contradictious, vacillating account of our general ideas, though he does not ask whether this very contradiction and vacillation may not be an authentic expression of untutored introspection. He then goes on to consider Berkeley's treatment of general terms as signs standing for clear-cut ideas which cannot be distinguished from images, and which are as incapable of being contradictiously comprehensive as they are incapable of being vaguely schematic. He points out, as others have, that Berkeley altogether failed to realize the essential 'woolliness' of our images, the sheer impossibility of making definite statements in regard to many of their characteristics. But he denies that such woolliness in an image has anything whatever to do with the comprehensive schematism of our conceptual references. He says also that Berkeley altogether failed to make plain what could be meant by saying that names acquire generality through 'equally representing', or 'indifferently standing for', a large number of particular ideas of a certain sort. Here, however, he seems to me to do Berkeley less than justice. For Berkeley *did* say that it was possible to consider a particular triangle merely *as* triangular, without attending to the particular qualities of the angles or the relations of the sides. And he also said that, in so far as we thus leave out of consideration much of what we perceive, what we say or prove will apply equally to an indefinite number of other particulars of the same sort, considered in the same abstract fashion. It seems to me that Berkeley has here directed attention to an extremely important phenomenological fact, the fact that things can come before us merely *as* this or *as* that, not only with certain

of their *parts* but with certain of their *sides* or *characters* in relief: this fact would certainly seem to have some important bearing on the facts of meaning and reference.

Holloway finds much to criticize in Hume's dispositional account of meaning, in his view that meaning lies in the *power* of words to evoke any one out of a vast number of resembling ideas which have been associated together through custom with the words in question. He says that this view would identify the meaning of a word with some *actual* associative series which it happens to arouse, regardless of the fact that there are *many* such series which it is capable of arousing, and which are all equally representative of its meaning. He thinks, too, that Hume's doctrine fails altogether to account for the fact that we sometimes *reject* some suggested similar brought before us by association as *not* properly comprehended in our meaning. We might, for example, reject a small porcelain swan which we happened to remember, as not properly to be regarded as a swan. Such self-criticism, he thinks, cannot be exercised by a mere associative series. Holloway thinks also that the tendency of our mind to move along *one* line of resembling associates rather than another, *e.g.* along a line of things resembling each other in colour rather than in shape, can only be understood if the distinction between different lines of resemblance is somehow 'intelligible in advance', and capable of being 'specified in advance'. I must say regretfully that I find these last contentions quite unintelligible.

Hume was, of course, hopelessly wrong in trying to *explain* association by resemblance through the workings of custom: this was, in effect, nothing but a mistaken attempt to reduce it to a case of association by contiguity. And he was also deeply wrong in thinking that he could in some manner *get behind* the fact that things resemble each other in a number of distinct ways, so that a series of things resembling each other closely in *one* way, may be quite unlike each other in *another* way. But there are parts of his work where Hume simply identifies relations (*e.g.* those of space and time) with 'manners or orders in which objects exist' or 'manners in which impressions appear to the mind', without erecting such 'manners' into substantive impressions, and without denying that there may be more than one 'manner' in which things may be arranged or appear. There would accordingly be nothing un-Humean in recognizing that things can resemble each other in a number of distinct 'manners', and that our minds can be therefore led on from one thing to another in a number of distinct directions. Hume was, in fact, acknowledging a plain fact when he held that our mind is, at varying times, a seat of varying dispositions, which draw it on along different lines of similarity, so that at one time it passes readily from a red cube to other red objects, and at another time from this same red cube to other cubical objects.

I can see no reason why such a readiness, once we concede it existence, should be bound up with a given, actual, particular sequence of ideas, as Mr. Holloway thinks it must be. Nor can I see any reason why a mind subject to such a readiness should need to specify its direction in advance, as Mr. Holloway seems to think it should be: to *have* a readiness doesn't mean that one must know about it, or need mention it. And a readiness for ideas and images of a certain sort is surely just something that *will* explain our ability to select from, and to discriminate among, our various images, and to accept some as appropriate while others are rejected as discrepant. For a man genuinely set or prepared for anything which resembles something else in a certain manner, will naturally hail certain

objects as relevant, and dismiss others as irrelevant, whether or not such objects are brought before him as the result of his own random reproductive processes. All this Hume plainly recognizes in that well-known passage (not quoted by Holloway) where he tells us how, when we talk about triangles, and say something only applicable to the particular sort of triangle represented before us, the other types of triangle immediately 'crowd in upon us', so as to make plain that what we are saying won't invariably accord with a triangular nature. Hume may, in fact, be said to have discovered 'determining tendencies' long before Ach distinguished them in 1905: his only error lay in failing to differentiate them from the reproductive tendencies which accompany them, and which he treated under the heading of 'custom'. Hence his wonderment at their operation, and his tendency to regard as an 'extraordinary circumstance' what is really an all-pervasive feature of our mental processes.

We may note, further, that Holloway's own theory of general meanings (which isn't developed very adequately) must necessarily be in terms of our readiness for certain sorts of speech and action, an account open to exactly the same objections (if indeed they *are* objections) that he makes to Hume's theory. Whatever the shortcomings of Hume's actual exposition, it is surely plain that if it is possible to work out a dispositional theory of meaning in terms of the kinds of *act* a man is ready for, it must be equally possible to work out a dispositional theory of meaning in terms of the kinds of image or idea that he is ready for. In neither case will it be necessary for us to be able to specify the direction of such a readiness in advance, in order that we may *have* the readiness in question.

From sensationalist analyses of meaning and reference Holloway passes on to consider analyses in terms of 'universals', whether these latter be regarded as really present in things, or merely resident in our heads. To this type of analysis he makes several very sound, as well as several novel, and (to my mind) rather fantastic objections. He says that to speak of a universal as an entity uniformly present in a large number of instances, is to give no explanation of the fact that things may exhibit some quality *in different degrees*, that one thing may, for instance, be redder than another, or more nearly a tumble than another, and so forth. It seems to me that no theorist of universals would be seriously bothered by this difficulty. For the 'realm of universals' is surely thought of as a sphere of entities bound together by many relations, and ordered in a large number of distinct dimensions: they may be closely or remotely like each other, they may be more or less readily compatible with each other, they may surpass or fall short of each other in a number of distinct ways. There is therefore no difficulty in saying that A will be redder than B whenever it instantiates a shade of redness of inferior intensity to the shade of redness instantiated by B. I should not regard this as an extremely illuminating way of speaking, but I cannot see that it leads to any special difficulties.

Holloway also professes to find great difficulties (for the theory of universals) in the relation between a universal and the general term which stands for it. For he thinks that a believer in universals must necessarily want a general term to stand for a universal in some 'genuinely thinking and cognitive manner', by which he means, apparently, that it mustn't merely *pop* up in the mind in a reflex fashion, when some universal happens to be considered. This leads him to maintain, further, that believers in universals must say that we never use general terms understandingly without also realizing *that* they stand for certain uni-

versals. And it leads him to maintain that believers in universals must also hold that we never make use of general names without *recognizing* that the verbal tokens we are employing *are* tokens of a certain general type or design. Holloway then points out that if this last recognition is itself to be formulated in words, we shall be landed in an infinite regress, since the tokens involved will have to be recognized as tokens of a certain general type standing for certain universals, and this recognition, if verbalized, will demand similar interpretative complexities, and so on indefinitely.

But why, we may ask, should a believer in universals be committed to this whole stupefying farrago? Why on earth should he suffer from such a silly horror of the 'reflex' that he doesn't think it possible to understand verbal tokens unless he is also conscious of them as tokens, and unless he realizes that they are of certain general types, and that they stand for certain universals? And why, if he must be thus self-conscious, must he formulate his self-consciousness in words, and then be committed to unending writhings of reflection? Surely the understanding use of names (whether standing for universals or anything else) demands only that we should use the name whenever we have its designatum 'in mind': it cannot require that we should also be reflectively aware of the name *qua* name, and of its customary connection with the designatum it stands for. This whole demand that there should be nothing merely habitual in our use of names represents, in effect, a denial of their conventional character: it is hard to see why believers in universals should be committed to such a primitive piece of silliness. That words are conventional expressions of our cognitive references, that their use on a given occasion is a consequence of unconscious social conditioning, doesn't at all mean that they can't be used understandingly, provided that the man who uses them, has their designata 'before him', whether in actual presence or anticipatory readiness. This is not, of course, to say that there may not be the gravest disadvantages in the whole treatment of general terms as names having universals as their peculiar designata. But if we decide to speak in this fashion, then it would not appear that Holloway has discovered any very desperate difficulties in the relation of general terms to the entities they stand for.

Holloway is on rather stronger ground when he goes on to demolish those theories of understanding which explain it in terms of universal mental *contents*, or concepts in the mind. He points out that it is hard to explain just *how* such contents mediate our verbal references to objects, concepts being in no sense images or likenesses of the things they enable us to refer to. He reiterates his previous, not very sound objection to all merely 'reflex' connexions between words and concepts: he thinks their connexion, to be genuinely explanatory, must be in some sense 'thinking and cognitive', and not merely automatic. He makes no attempt to deny the existence of all the experiences of *unanschauliches Wissen* which have been vouched for by careful self-observers, but he points out that such experiences admittedly vary both in form and intensity, and are in no sense necessary conditions of our understanding use of verbal signs. He objects further that if understanding consists merely in having the non-imaginal experiences in question, a man might conceivably acquire an understanding telepathically, without any previous acquaintance with the subject matter it concerned.

In this part of his work Holloway simply ignores various important theories which connect universality with conceptual references, without

making it pertain to any actual *content* in our minds: Ockham's obscure doctrine of mental acts as 'natural signs' or modes of signifying objects, as well as Brentano's doctrine of 'intentionality', should certainly have been considered. Quite obviously we can't throw light upon our verbal references to objects by connecting them with changes of an utterly different sort located in a wholly alien medium. The relation of such 'mental contents' or 'intelligible species' to the things we mean and speak of, is then made empirical and contingent. To be of use, the term 'concept' might stand simply for the phenomenological fact that objects come before us with that curious high-lighting of aspects, on which Berkeley commented. Or for the still stranger, more exciting fact that, in the total absence of appearing objects, there can none the less be something that we can only call a 'reaching out' towards their presentation, that objects may, in Humean phrase, be ready to 'crowd in on us' and that we are in the strange position of being able to know and say what things are thus ready to crowd in on us.

Holloway is right in saying that the laboratory findings in regard to imageless thought are wholly uninteresting and unilluminating, if all they reveal are certain highly variable, cobwebby affections, which some people have at some times, and other people never have at all, and which arise casually out of our intercourse with objects. They are important only if we frankly recognize the wholly non-existential *dispositional* character of whatever can be thus described as 'imageless', if we think of it as representing a trend, a drift, or an unrealized potentiality rather than as anything actual. We should, however, recognize all this, without seeking to minimize the importance of what is thus taken to be dispositional, or without seeking to deny the extraordinary fact that a man may be able directly to report upon his own momentary dispositions without needing to infer their presence from anything else. It has long been plain that one can't make sense of man's outward behaviour while suffering from a Megarian horror of the merely dispositional: this is even plainer in regard to the substance of his inner life. We may note, in passing, that Holloway's shock at the notion of an understanding acquired by telepathy is a little exaggerated: it is merely our misfortune that we have to learn everything the long way, and that it is impossible for us to become inwardly prepared for a sort of thing of which we have not previously had concrete experience.

In the remainder of his book Holloway is engaged in interesting analyses of our sign-using and intelligent behaviour, with which one can have little cause to quarrel. He gives an account of our understanding of signs very much on the lines of Morris in his *Foundations of the Theory of Signs* and in his *Signs, Language and Behaviour*. He says that one interprets a sign as standing for something, when one reacts to it (or is ready to react to it) more or less as one would react to the thing it stands for: interpretation is as Morris says, a mediated taking account of something. Our interpretation of words is not at all different from our interpretation of other signs: no mysterious leap occurs when we begin to understand verbal symbols. Holloway makes the obvious point that we needn't be conscious of our symbols, nor of the way in which they function as symbols, in order that we may use them significantly: it is only odd that he should imagine that many would deny this. He then proceeds to analyse intelligence in general, and criticizes many of the positions of the Watsonian reflexology, positions so worried to death in American controversy as scarcely to deserve further criticism. He notes that intelligent performances re-

present deviations, sometimes sudden deviations, from merely routine responses, but he points out that there are senses of 'habit' in which being intelligent is not at all incompatible with being habitual. Many of our habits are, in fact, flexible dispositions whose adjustment to the needs of the situation leaves room for intelligence. And it would be impossible to react intelligently to situations if one merely cut adrift from the precedents that have gone to the moulding of one's habits.

Holloway is plainly right in refusing to give any single, simple explanation of our intelligent adjustments: they have no unique initiating conditions, they may have many varied standing conditions, they are expressions of individual capacities which stand in some unknown relation to organic structure. He is concerned, above all, to demolish those theories which erect some inner *experience* of understanding either into the essence, or at least into an indispensable initiating condition of our intelligent performances. He points out that we can't define intelligence by means of the inspiring presence of an inner act of understanding, without thereby making it impossible to predicate intelligence or unintelligence of this inner act itself. And yet everyone would agree that it is as possible for a man to be intelligent or stupid in his thoughts as it is for him to be intelligent or stupid in his actions. Holloway also points out that ballerinas, pugilists, singers and other intelligent performers, very seldom think before they act, and that they would probably be 'put off their stroke' if they attempted to do so. They can be said to think with their larynxes, feet, fists and other organs, rather than with their heads. He also says that, while some of our intelligent performances may be surrounded with a great glow of felt illumination, this is wholly absent in other cases, and that such a glow is, in any case, a highly deceptive criterion of intelligence, since one may feel extraordinarily illuminated and competent when one is about to behave with abject incompetence and stupidity. He regards experiences of illumination as *indirect* criteria of intelligence, which merely suggest that we can solve a variety of problems: our actual success in solving such problems is the direct performance criterion of intelligence. He thinks that it is owing to a linguistic confusion that men have been led to make of inner experiences of understanding, realization, etc. a *sine qua non* of intelligent action: they are confused because the same words 'understanding', 'realization', etc. stand *both* for an inconclusive, unreliable experiential criterion of intelligence and for a conclusive, reliable performance criterion.

One must acknowledge the justice of most of these contentions. We speak of intelligence or cleverness when a man really can solve problems and get round difficulties, and not when he merely *feels* clever. And it is possible to say that a man has said or done an extremely intelligent thing, even though he was barely aware of what he was saying or doing. So that the path lies open for extensions of the use of 'intelligence' to cover the reflex 'wisdom of the body' or the startling performances of electronic brains. But it does seem to me that Mr. Holloway has allowed himself to be overcome by a quite unwarranted, merely fashionable preference for the performances that are overt, as opposed to those that take place in our heads. Even if experiences of illumination are highly deceptive criteria of intelligence, the successful solving of a large number of problems in one's head may be a very reliable criterion of intelligence. The mere fact that the same problems could be solved on paper is not at all to the point. We must not abandon the obvious truth that inventiveness, penetration

and ingenuity may as much be shown in the manipulation of our ideas as in the contraction of our muscles.

We may note, further, that Holloway's argument in regard to the performances of ballerinas, pugilists and other adepts is infected with a serious ambiguity: he forgets that the feet, fists, larynxes, etc. of these gifted performers are not merely external physical objects, but have an interior or felt side. A ballerina operates with seen and felt feet, a singer with a heard and felt larynx, and there is an important sense, horribly out of fashion at the present moment, in which even the performances of these *virtuosi* may be said to take place in their minds or heads, *as well as* in their bodily organs. One might even go further, and say this much in favour of the traditional connexion between intelligence and 'what takes place in our heads', that by far the most intelligent performances of which we are capable are those which take place *predominantly* in our heads. For it is in our heads that we can simultaneously prepare ourselves for a large number of complex, alternative possibilities, can develop them in various truncated and economical forms, and can then arrive at some solution or decision which does justice to them all. Hands, tongues, pens, limbs and so forth merely lead up to, and take their lead from, such supremely intelligent performances.

The last four chapters of Holloway's book deal with verbal language: he points out that our use of spoken symbols always has a supporting context of gestures, intonations, timing and so forth, which are not brought out in written speech. He points out also that the uses of language are infinitely various, and that it is merely a philosopher's exaggeration to imagine that the main function of speaking is to make statements capable of being true or false. Here it seems to me that Holloway is himself exaggerating: commands may predominate in the talk of strong, silent men, and promises in the discourse of Oxford philosophers (I hardly ever make them myself), but no one who has listened in on some conversation in a bus can doubt that the descriptive and narrative function bulks overwhelmingly in human discourse. Holloway says a great deal that is true and interesting about the uncertain, unsettled character of everyday usage, and of its total unlikeness to the regimented precision of symbolic calculi. He shows, too, how the attempts of philosophers to impose rigid exactness on this elastic material will necessarily result in a series of linguistic strains which will drive the wretched philosopher on from one formally exact system of speech to another. At last he will become wise, and will realize that any attempt at an exhaustive systematization of ordinary diction is foredoomed to failure. One is here reminded of much that was said by Hegel in regard to the freezing, stereotyping function of the Understanding, as well as the dissolving function of Dialectic. One only misses the suggestion that there might be some philosophically *reasonable* form (or forms) of speech which might achieve some of the merits, without copying the defects, of ordinary language. One may say, in conclusion, that Holloway has certainly written an extremely interesting and intelligent book, which stimulates one most deeply even when one disagrees with it most.

J. N. FINDLAY.

Immanuel Kant. Ontologie und Wissenschaftstheorie. By GOTTFRIED MARTIN. Kölner Universitätsverlag, 1951. Pp. 244.

THIS is a very good book, It is informative, lucid, and short, it does not suffer from a welter of footnotes, and the references given to the

works of Kant and other philosophers are adequate without being oppressive. Professor Martin's method of exposition is one which deserves careful study in this country since we still suffer from a tendency to treat Kant as the lineal successor to Locke, Berkeley, and Hume and to forget the extent to which his thought was moulded by influences of quite a different kind.

Professor Martin's contention is stated in his opening sentence. It is that 'the *Critique of Pure Reason* is a sea fed by two great rivers. One of these is the new science of nature and the other is the old ontology.' His purpose is to show the way in which these streams mingle in Kant's completed view, but this does not mean that he subscribes to what has been rudely described as 'the General Election theory of the Critique'. He is not maintaining that there was a party called the rationalists and another called the empiricists and that Kant discovered a helpful formula about synthetic *a priori* propositions which enabled these parties to patch up their differences and form a coalition government. Professor Martin's view is rather that Kant understood the complexities of words like 'unity' and 'reality' as Plato, Aristotle, and Aquinas had done and recognised the fact that the philosophical Levellers of the *Aufklärung* were creating insoluble difficulties for themselves by ignoring these complexities. Hence the first part of his book is an exposition of Kant's use of '*Einheit*', while the second deals with '*Sein*'.

As far as 'unity' is concerned, there are three different types which are dealt with in the *Critique*, namely that of things as spatio-temporal (in the Aesthetic), that of the world or the whole of what there is (in the Antinomies) and that of nature as a system of laws (in the Analytic). Professor Martin discusses these in detail with reference both to the controversies between Newtonians and Leibnizians as revealed especially in the Clarke-Leibniz correspondence and to modern developments in Geometry, Logic and Physics. What he has to say here is interesting and important. It has not been said before in any book on Kant, though the general view of axiomatic systems and of the reason why debates of the type expounded in the Antinomies are futile was understood by Moritz Schlick and the Vienna Circle. As a matter of history one may legitimately doubt whether Kant recognised as clearly as Professor Martin suggests that he did the force and the implications of his own analysis, but this is not a very important point. In any case Professor Martin agrees that Kant would probably have been rather shocked by the discovery of antinomies or paradoxes in the sacred domains of logic and mathematics, though this in fact finally confirms his rejection of the Leibnizian doctrine that at any rate some relational statements are respectable because the thoughts which they express are the thoughts of God and not the thoughts of men. To state the case very shortly, we might express Kant's contention that the mind (sensibility, reason, and understanding) imposes unity on nature by saying that we create models in terms of which we explain to ourselves what goes on. This is not a psychological hypothesis, it is just a description of the way in which we come to discover truths about the world; and there is no other way in which creatures like ourselves can make such discoveries. Where Kant went wrong was in supposing that the models created by Aristotle, Euclid, and Newton were the only possible models, and mathematical logic, non-Euclidean geometry and quantum physics prove that he was mistaken. But Professor Martin emphasises that this is not the end of the story. These models do seem to be in some not very clearly explained sense fundamental. They are

not *a priori* in Kant's sense but there is something in them which at least calls for further enquiry and analysis.

The difficulty now is to produce an account of reality which is consistent both with Kant's doctrine, especially in the second and third *Critiques*, and with this analysis of 'unity' in the *Critique of Pure Reason*. The obvious line to take and it has been taken by many commentators, is that the *Noumenon* or *Ding an sich* is a pre-Critical survival. If he is to be consistent Kant has to say that 'reality' means verifiable location in a space-time continuum for he has demonstrated that statements which purport to refer to non-temporal existents are strictly meaningless. But in fact this was not Kant's view. He manifestly held that significant statements could be made about God, Immortality and above all, Freedom; and since he claimed that the two props of the Critical Philosophy were the transcendental ideality of space and time and the objective reality of freedom, he clearly did not believe that the second of these was an empty phrase. But what did he mean? Professor Martin shows conclusively that in the ordinary use of 'analogy' it will not do to say that statements about God are analogical. There is no analogy between divine and human understanding or between divine and human will. He holds, however, that in Kant's view and as a matter of fact human spontaneity is a different type of reality from other occurrences and it is something which we are conscious of (*bewusst*) though it makes no sense to describe it as something we know (*erkennen*). I confess that I do not follow this argument, and I wish that on this very difficult point Professor Martin had stated the case at rather greater length (an unusual criticism of writers on Kant's account of noumena!). My difficulty is this. I can distinguish perfectly well between 'I did it on purpose', and 'I did it by accident', or 'from habit'. A man who suffers from schizophrenia does not act spontaneously in the sense in which a normal man does, and so on. But if it is argued that spontaneous acts are in some sense 'noumenal', does this mean that they cannot be an object of scientific study? Certainly I cannot observe a choice, whether my own or someone else's, in the way in which I can observe a change of colour or of temperature. The grammar of 'choice' is different from that of 'cause'. But I cannot at the moment see that any momentous ontological conclusions can be drawn from this. I am inclined to agree with Professor Martin that Kant believed himself to have made an important discovery here and to have formulated an argument for the existence of God and the immortality of the soul which practical but not theoretical reason would support, but I also think that this was one of Kant's great mistakes.

This point, however, is really outside Professor Martin's terms of reference and it would be unprofitable to press it further here. He has made as good a case as can be made for Kant's attempted distinction between the sensible and the supersensible and for the odd doctrine that there are entities which are both real and unknowable—and that is what he set out to do.

T. D. WELDON.

The Vindication of Metaphysics. A Study in the Philosophy of Spinoza. By RUTH LYDIA SAW. London: Macmillan & Co. Ltd., 1951. Pp. 173. 12s. 6d.

THE title and sub-title together accurately describe the double ambition of this book of 170 pages. In a short introductory chapter Miss Saw tries

to clear the ground of those anti-metaphysical doctrines or prejudices which may prevent contemporary philosophers from seriously attending to Spinoza's arguments. One may doubt whether any such general vindication is either possible or necessary, and whether general arguments about the 'function' and 'job' of philosophy can prove anything on either side; Spinoza was not a man with a job or a function which was also Hume's or Kant's, and there is no factory of truth in which philosophers have an allotted place at the bench. Spinoza produced a set of specific arguments to support his conclusions, and anyone who would show that his conclusions cannot be true must show in detail that his reasoning was bad reasoning or that his premisses are untenable. And those who have thought that they could *prove, a priori*, and without further examination of the text, that Spinoza had attempted the impossible, will not be converted by this introductory chapter; for Miss Saw wrongly identifies her target as 'the verification theory of truth' (p. 5) and as 'a definition of truth' (p. 6); the verification principle was not, and could not be, a theory or definition of truth, but (at the best) a restrictive criterion of literal significance.

The exposition of Spinoza which follows the introductory chapter is informally written and makes no claim to be complete or to follow the text with care and scholarship in the manner of H. H. Joachim's *Study of the Ethics;* there are comparatively few quotations and only the broader issues are discussed; in fact this is not a text-book but a set of reflections on Spinoza's themes. But even the student who wishes the *Ethics* to be studied for its own sake, and without apology, must be grateful that at least the themes are mentioned and some interpretation provided.

Miss Saw does not follow the order of exposition of the *Ethics* itself; chapter I is of Knowledge, Truth and Error, which suggests that Spinoza's metaphysics is here to be explained as following upon his logic; and this is no less, and no more, legitimate in the case of Spinoza than in the case of Leibniz; for the possibility of intellectual error, and the discovery of a method of eliminating the possibility, can be taken as the starting point of rationalist philosophy. Miss Saw emphasises (p. 25) that 'the important contrast [in Spinoza's logic] is between complete and incomplete knowledge, not between truth and falsity'. To fall into error is to be the victim of fragmentary or incomplete knowledge; logical interconnectedness (p. 27) is the mark of genuine knowledge, and interconnectedness, and therefore knowledge, admits of difference of degree. Unfortunately in explaining (p. 28) knowledge of the second level, and the 'common notions' on which it depends, Miss Saw writes of Extension and Thought as 'properties' and supposes common notions to be notions of these properties; but propositions 37, 38 and 39 of Part II, taken together with the reference made to Lemma 2, make it clear that common notions are those axioms of mathematical physics which are universally applicable to Nature—'such for instance, as motion and rest, and their laws and rules' (Tractatus, *Theologico Politicus*, chapter 7). Chapter II, on Spinoza's method, unfortunately does not consistently follow and criticise his own very explicit account of his method and of his decisive opposition to Descartes in respect of method; and there is a continuing confusion (pp. 42-45) between logically necessary propositions and definitions; there are sentences which are more difficult than Spinoza's, *e.g.* (p. 47)—'If such a state of affairs is thought of as appropriately expressed in a series of definitions, each step would have to be names, but the thing named would have to be seen as the ground for the next step; that is to say, it would not

be a proper name'. On p. 53—'Spinoza never refers explicitly to the realist-nominalist controversy, but I think it is clear that he would have been a nominalist at the deepest level'—but he does refer in *Ethics*, part II, prop. 40, also in prop. 17 and prop. 18, scholium I, and in respect of all class terms he shows himself a blunt nominalist following Hobbes; and the nominalism is essential to his metaphysics. Chapter III, on God's necessary existence, discusses the ontological argument without (I think) sufficiently underlining the identification of God, as the unique self-determining substance, with nature, taken as the total system of things; it is this identification which gives his ontological argument, or the equivalent of it, its peculiar form; beneath the forms of an ontological proof there lies the argument that, if there are any created things and if their coming into existence is to be understood, there must be one all-inclusive, self-creating cause of which they are finite modes or particular manifestations; any other account of creation either involves an abandonment of rational explanation at some point or leads to contradiction. But 'an empirical test' (p. 73) for belief in God's existence is altogether remote from Spinoza's intentions. Chapter IV, of God, Substance, Attributes and Modes, is less than ten pages, and explanation of these doctrines really requires more space; and the doctrines are not always presented faithfully, *e.g.* it is probable that the correlate under the attribute of Thought of 'facies totius Universi' under the attribute of Extension is the 'infinita idea Dei', which cannot be rendered (p. 81) as 'the whole system of thinking of human beings'. Chapter V, on God, suffers from Miss Saw's reluctance to accept Spinoza's peculiar use of this word, which separates him so widely from Christian theology (p. 94): 'the fact of the matter is that Spinoza ought either to allow himself to speak of God and men, or substance and its modifications, and not confuse the concepts which belong to different contexts'. This is a refusal to enter into the design and motive of his philosophy, a refusal which reflects itself throughout the chapter in efforts to make Spinoza nearer to common-sense Christianity: it is surely impossible to claim (p. 98) that 'definite substance turns out to be a person', when Spinoza is constantly protesting against the imagination of a personal God. In chapters VI, VII and VIII, on the World, Men and Moral Theory, Christian apologetics are liable to replace direct dealing with Spinoza's argument, and particularly in the conclusion (p. 158) 'Spinoza has produced the characteristics which a God must possess in order to be effective, but he has had in mind the person whom his description fits'. The final chapter, Appraisal, presents metaphysical beliefs (p. 162) as 'general presuppositions' connected with 'a satisfactory emotional life' and as 'beliefs which are justified by appropriateness of feeling and of action'; such a vindication of metaphysics can have no relevance to Spinoza, for whom it was a first principle that philosophical truth must never be adapted to conform to human need and feeling.

Certainly this book contains some observations on Spinoza's philosophy, and perhaps also on metaphysics as an inquiry, which may interest the general reader; I have been compelled to draw attention only to the points which seem definitely misleading.

Misprints are on p. 56: 'idea dubio' for 'idea dubia': p. 80 'facies totius universiae' for 'facies totius universi'.

STUART HAMPSHIRE

Gestalt Psychology. By DAVID KATZ, Ph.D. Translated by Robert Tyson, Ph.D. London: Methuen, 1951. 12s. 6d.

GESTALT Psychology has been something of a cult, its followers making exaggerated claims for the applicability of its main principles. Professor Katz has been on the fringes of the movement rather than a devotee. His first work, on colour phenomena, which was in the spirit of the Gestalt movement, appeared in 1911 before that movement really began. In this work, a translation of the second German edition, he gives an appreciation and critique of the Gestalt Theory in most of its aspects, and adds an account of some work of his own which has been guided by the Gestalt point of view. Whilst acknowledging the fruitfulness of that point of view, he is always ready to admit its limitations, and the book is welcome for this reason alone.

He stresses rightly that Gestalt Psychology cannot be understood except as a reaction to the psychology which existed before. I do not think, however, that he characterises that reaction adequately. He continually refers to Gestalt Psychology as 'anti-atomistic' or 'holistic', but he does not seem to have considered sufficiently with what justification these epithets could be applied to a scientific theory. This is born out by what he has to say of the relations between Gestalt Psychology and Phenomenology. Reliance upon the latter led to a denial of the view that what we perceive is sensations, a view which was adopted by the older psychologists on epistemological grounds. This accounts for the platitudinous nature of some Gestaltist 'discoveries'. (*Cf.* "The five laws listed are convincing evidence that environmental objects are seen exactly as they appear at the time of observation," p. 27.) Gestalt Psychology is holistic both in the sense that it involves a denial of atomistic epistemology and because it emphasises the role of context as a function of behaviour; the notion of a 'gestalt' has to play both parts. Professor Katz stresses the experimental work of the Gestaltists and does not introduce the question of the meaning of the term 'gestalt' until page 91. This may be due to some realisation of the confusion. Certainly he is aware of other confusions with epistemological issues, as, for example, in the Gestaltist attempt to account psychologically for the making of correct inferences.

In his survey of Gestaltist work, he considers first the emphasis upon perception and continues by considering their work in most psychological fields, concluding with an estimate of the future of the movement. As a survey, the book falls between two stools. It is too brief and on occasions too sweeping for an authoritative work (the chapters on medical, comparative and social psychology are so brief as to be almost valueless). As an introduction, it does not make Gestalt Psychology any less puzzling than other works, and may even cause bewilderment due to the sequence of the argument and the tendency to use abstruse terminology. The bibliography is curiously incomplete, there being no references to Wertheimer or any of Kohler's later work in English. The effects of the movement on American Psychology is generally ignored.

This is not a work which will provide a final evaluation of Gestalt Psychology. It does not sort out what is philosophical from what is not. It is nevertheless welcome, as coming from one who is neither exclusively for, nor exclusively against, the Gestalt Theory. It thus provides something of a balanced point of view despite its limitations.

DAVID HAMLYN.

A History of the Theories of Aether and Electricity. Vol. I. The Classical Theories. By SIR EDMUND WHITTAKER, F.R.S. London: Nelson, 1951. Pp. xiv + 434. 32s. 6d.

PROFESSOR WHITTAKER'S standard history, covering the period up to the close of the nineteenth century, was published originally in 1910. Almost at once a supplement began to be needed, for the advances in physical theory which began in 1905 (Einstein's *annus mirabilis*) soon left far behind the ideas whose development he had so absorbingly described. Since that time physicists have had difficulty even in keeping afloat on the tide of new theories and discoveries: the historian's Olympian position has been quite inaccessible. But now at last it is again possible to take breath and to assess the changes that have come over physics, and we are lucky in having Professor Whittaker once more as our guide.

The volume under review is the first of a pair, of which the other will be called *The Modern Theories.* It begins effectively with Descartes in optics and with Gilbert in magnetism and electricity, and closes around 1900-10. (Relativity and quantum theory are here mentioned only in passing, but they will be the main topics of Vol. II.) The treatment is in no sense 'popular': to read it in full a fairly advanced knowledge of mathematics is required. But there are no formulae to speak of before page 61, and any reader of MIND will find things of interest in the first two chapters, especially for an understanding of seventeenth and eighteenth century philosophy: Descartes, Leibniz, and Berkeley never distinguished between their epistemological and physical enquiries as sharply as we do. Light is also thrown on philosophical disputes by Berzelius' discussion (1812: *cf.* pp. 79-80) 'whether the electricities and caloric are matter or merely phenomena'; for he makes the answer to this question turn on the answer to the logical question, whether 'ponderability' is or is not to be made a necessary condition of 'materiality'.

Quite apart from points of direct philosophical interest, this is a useful source-book for the logic of the theoretical sciences. Professor Whittaker's manner of exposition is strict, and illustrates clearly the methods of modern physical theorising. From it one can see how great a part of the aim of physics has been to correct and elaborate our *conception* of such things as light, how negligible a part to *generalise* about them. Here there is neither the busy 'accumulation of confirming instances' one reads of in the logic-books, nor continual speculation about comparative degrees of probability. All is 'pictures', 'models', and 'representations'; and truth is measured not by the number of facts inventoried, but by the faithfulness of the representations developed (N.B.—the quotation from Faraday on p. 172). Discovery, cause, conjecture, intuition: the place such notions have in physics is well exemplified by the use Professor Whittaker makes of them, and it departs at every point from that implied in the standard logical treatments of the subject. Logicians, it seems, have yet to learn that physics is not 'the natural history of the inanimate'.

Two other things are worth noticing: (i) how much less sharp than has often been supposed are the differences between the wave and particle views of light (17n., 24n., 317, 358); (ii) the extent to which physicists advocating 'field' theories, including Whittaker himself, see themselves as opposed to action-at-a-distance theories, and *vice versa* (59, 172 n. 5), though to the outsider the notion of a 'field' may look like a way of reconciling oneself to action-at-a-distance, rather than opposing it. A clarification of this second point would elucidate also the rather confused

remarks (31-2) about 'unobservable interphenomena' being 'assumed to exist in order to set up continuity between separated events which are in fact observed'. An important point is hinted at here; but it is misleading to mark the *logical* contrast between theoretical entities and laboratory phenomena by the use of such phrases as 'unobservable', '*in fact* observed'.

One last matter, but an important one: at the outset, Professor Whittaker divides philosophers into two classes, 'those who have speculated in the grand manner (λογικῶς)' and those who more modestly 'proceed to their general philosophy by extending their conceptual analysis of natural science (φυσικῶς)'. As he says, this way of dividing philosophy is of long standing. But it enshrines none the less hoary a fallacy for all that. For this division has itself been a perpetual source of philosophical trouble, suggesting as it does that the only way of tackling problems methodically is to employ the techniques of physics. Surely two of the chief lessons of modern philosophy are, that to proceed λογικῶς it is not necessary to be grandiose, and that the natural sciences need not be left with a monopoly of modest intellectual constructiveness.

STEPHEN TOULMIN.

Les Chemins du Savoir dans les Derniers Dialogues de Platon. By PAUL KUCHARSKI. Presses Universitaires de France, 1949. Pp. 404. 1,000 fr.

PLATO'S view of the tasks and techniques proper to philosophical enquiries changed substantially in the course of his writings. M. Kucharski insists that the change is to be construed primarily as a revolution in methods, and only secondarily as a change of theory. This is a move in the right direction, but not far enough. For M. Kucharski sets up as critic not only of Plato's tactics but of his plans of campaign. And while he gives thought to some (though not all) of the technical innovations in the later dialogues, he gives little to the new problems in the elucidation of which Plato shaped and tried his tools. He distinguishes ably, if over-sharply, various patterns of διαίρεσις and συναγωγή which are used and annotated in those dialogues; the articulation of a genus into its limiting species and the extraction of one species by dichotomous division are held apart and traced respectively to the "Hippocratic method" and the rules for "dialectic" which turn up side by side in the *Phaedrus*. But he tends to confine Plato's key-problems, from the *Phaedrus* onward, to those which are—in M. Kucharski's view—solved or else engendered by the use of these favoured methods. One consequence is an odd inversion of the argument concerning existence, non-existence and negation in the *Sophist*.

Similarly, by neglecting the succession and transformation of the questions which control Plato's earlier dialogues, he oversimplifies the contrast between the later dialectic and the old Theory of Forms. This theory, as it appears in the *Phaedo* and *Meno* and subsequently under fire in the *Parmenides* and *Sophist*, is described as "un tout singulièrement cohérent". But he fails to substantiate the claim. He maintains that "l'adéquation entre les noms et les choses en soi est un postulat implicitement reconnu dans la doctrine de l'εἶδος. A chaque nom répond une Forme définissable"; and for no very clear reason he regards the suppression of this postulate in the later dialectic as part of a process characterised as "getting beyond language". But he admits elsewhere that in a

résumé of the old theory in the *Parmenides* Socrates refuses to correlate a Form with every "common name". And he takes this refusal to show that the recognition of a Form involved a value-judgment; noting that, by contrast, the dialectic practised in the *Sophist* and *Statesman* is said by Plato to carry no such evaluation of its subject-matter. But he confuses this distinction in turn by connecting the later methods with attempts to construct various value-hierarchies and concluding that "aux yeux de Platon il n'y a point de différence entre une science 'normative' et une science 'objective'". (This argument collapses independently. It pretends that the inventories of motions in the *Laws* and of "goods" in the *Philebus* are systematic divisions of a genus.)

He further argues that, whereas the old theory was directly concerned only with the relation between Forms and their individual instances, the new dialectic enquired only into the interrelations of genera and species. But he fails to bring the *Meno* under this rubric, and a reading of *Protagoras* 329-331 might have convinced him that at an earlier stage the logical connexions of genera and species were not discounted but radically misconstrued. What is more important, it is wholly arbitrary to force the εἴδη of the later dialogues into the genus-species mould when this means obliterating such cardinal distinctions as that summarised in *Sophist*, 253c 1-3.

It was a function of Plato's later dialectic (but one to which, judging from his remarks, he attached decreasing importance) to criticise and implement the vocabulary with which it operated. Here again M. Kucharski detects Plato's master-plan for getting "beyond language". So he misses or misconstrues Plato's new interest in linguistic problems—problems not of vocabulary (as this is discussed in the *Cratylus*) but of syntax, the *ἀγήρων πάθος λόγων* which survives translation (*Philebus*, 15d-16a). Consequently he does not grasp the effect on Plato's epistemology of the demonstration, culminating in the *Sophist*, that a name or set of names can be used to say nothing true or false.

There is a bibliography, not more remarkable for its omissions than for some of its inclusions.

G. E. L. Owen.

The Conditions of Knowing. By Angus Sinclair. London: Routledge and Kegan Paul Ltd., 1951. Pp. 260. 21s.

This book gives the impression, rather like *Language Truth and Logic*, of offering a *solvitur ambulando* for every philosophical question. But instead of equipping us with new techniques, it proposes that we should adopt a new theory (or, rather, that we should recognise a theory which many of us are said to hold already without knowing it) in terms of which the old problems disappear. The theory is that all our thinking is conditioned by certain "attitudes" which in turn are conditioned by others of a higher order and finally by one (or more) of the highest generality. The ultimate "attitude" is a kind of epistemological theory. The main thesis of the book is that there is such a theory embedded in Western culture; that it is misleading and tends to perpetuate "confusion and ineffectiveness" in all fields from theoretical physics to political activity; that a change to a better "theory" is already taking place. In particular, it urges us to do away with all philosophical sub-theories dependent on the old one; each of the chapters on "Progress", "Order", "Time",

"Space", and so on, argues that all classical and most contemporary discussions of its titular topic are vitiated in a characteristic way by involvement in such a sub-theory.

Mr. Sinclair's task for epistemology looks very like Collingwood's task for metaphysics, except that the absolute presuppositions are to be corrected as well as studied. The offending theory cannot be formulated, but can only be indicated as something that accounts for the family likenesses among a set of false philosophical assumptions; for instance, that there is absolute time; that there are Facts; substances and qualities; knower and known; universals and particulars; and so on. It is not a fact that there are sensa, facts, absolute space, etc.: it is a theory that there are, conditioned, like all our theories, by our own higher-level epistemological "theories".

The author's hardest job is to explain what these "theories" or "attitudes" are. They are "ways of selecting and grouping in attention", which are characteristic of, indeed constitutive of, all our experience. I experience things as spatial, events as causally related, society as progressive, because I perform certain selectings and groupings and not others. The rubric is "When I say that a situation is orderly"—or spatial, temporal, causal, factual, etc.—"I am not saying anything about that from which the selection is made (which cannot meaningfully be said to be either orderly or disorderly) but am saying something about the selection". This commits the author to explaining (i) "that from which the selection is made", and this he identifies with "ultimate reality"; (ii) the principle of selection; this is a "criterion" appearing under several aspects such as "relative simplicity" and "permitting forecasting" but most prominently—despite a *caveat* against an imputation of pragmatism—"unmanageable consequences", "the penalty of confusion", and the like (ch. xxi).

Mr. Sinclair has a short way with opponents; his technique is "The objection may be made that p: but p is a consequence of the theory I am rejecting". Such rebuttal, however valid, is seldom cogent. Mr. Sinclair will need bigger guns to defend the Bradleian doctrine that reality is co-extensive with experience (p. 155); or the view that "knowing" (sometimes equated with "experiencing") is an activity, no reference being made to dispositions (p. 36); that there is no such subject as logic, but only a historical study of what has passed as such (p. 194); or that there are no statements, only men using "series of linguistic contrivances" (p. 188) for the purpose of "causing other men to adopt certain attitudes" (p. 114). And there are serious difficulties in the basic "attitude"—"theory"—"way of selecting" thesis: in what sense, for instance, do attitudes "determine" or "lay down limits within which must fall" opinions and actions? Mr. Sinclair's example is that if we know X has the "scientific outlook" we know that X's opinions will be formed with strict regard for the evidence. But how do we know what X's outlook is except by noting, among other things, how he forms his opinions? And in the wider field, Mr. Sinclair would agree that we don't first have high-level epistemological theories and then come by the notion of cause; but is there even a logical priority? Even if there is such a process as "following ways of selecting and grouping in attention", it is courting the intellectualistic fallacy to say that the process is carried on by reference to a criterion which is perhaps never recognised. No doubt it is a fact that alternative "ways" would produce "unmanageability"; but it is a fact which needs very careful manipulation if we are to avoid the paradoxes

of a thorough-going pragmatism. In short, there is no simultaneous answer to the Kantian question "What are the conditions of our experiencing things as we do?" and the pragmatist question "What is the justification for our believing what we do?"

There is much that will be welcome even to those who do not share the author's general theory: the insistence on the non-elementary nature of sense-data, for instance, or the highly original and interesting chapter on Language and Statement. But to mention the merits of the book would be, in Mr. Sinclair's view, to accept conversion to his epistemology; and this must be my excuse for emphasising the obstacles to conversion. The book is attractively produced and free of misprints, but it contains a diagram (p. 142), illustrating stereoscopic illusion, which doesn't seem to work. One feels these things are better done in psychological text-books.

BERNARD MAYO.

The Explanation of Human Behaviour. By F. V. SMITH. London: Constable, 1951. Pp. 267. 18s.

THE first part of this book attempts to define what contemporary psychologists regard as a satisfactory method of explaining human conduct; the second part illustrates the manner in which six psychologists exemplify this general method. Each of these—McDougall, Lewin, G. W. Allport, J. B. Watson, Clark Hull and Tolman—receives a separate chapter. The book is a good text-book giving a more detailed and critical description of the principal systems in contemporary experimental psychology than any provided hitherto. The best sections are those on McDougall (showing the difficulty of giving any precision to any of his concepts and the lack of empirical evidence for his theory of instinct) and on Clark Hull (a useful guide to the tough writings of the original).

Smith shows that, underlying their differences, the six psychologists are influenced by a common conception of their task as the search for causal explanations and that this is a very naïve conception derived from 'patterns of procedure' characteristic of early phases in the history of the physical sciences. Smith remarks that some of our commonsense explanations of human behaviour are no less rigorous than those offered by the professional psychologist and suggests that Freud may have pointed the way to a more satisfactory psychological method of explanation. Unfortunately this topic is merely raised as a postscript and not discussed.

The book has merit; it shows how naïve and confused psychologists are about the logical and philosophical problems which concern their subject; it not only gives a summary of what has been said by the psychologists but also makes explicit their assumptions and their use of technical terms; if it closes at the point where the discussion seemed to be raising the most interesting question (Freud or Clark Hull?) it must be remembered that the function of the book is to serve as a text-book. Those philosophers who are interested in empirical psychology will find it a useful one.

ROBERT THOMSON.

Received also:—

E. Altschul, *Die Lehre vom Bewusstsein*, Salzburg, Im Selbstverlag des Verfassers, 1951, pp. 32.

A. G. Alvarez, *Introduccion a la Metafisica*, Mendoza, Universidad Nacional de Guyo, 1951, pp. 393.

Aristotle's Syllogistic from the standpoint of Modern Formal Logic, by Jan Łukasiewicz, Oxford at the Clarendon Press (London: Geoffrey Cumberlege), 1951, pp. x + 141, 15s.

I. M. Bochenski, *Ancient Formal Logic* (Studies in Logic and the Foundations of Mathematics), Amsterdam, North-Holland Publishing Co., 1951, pp. vi + 122, 12 fl.

L. Brunschvicg, *Ecrits Philosophiques*, Vol. I, L'Humanisme de l'Occident —Descartes—Spinoza—Kant, Paris, Presses Universitaires de France, 1951, pp. 319, 1000 fl.

H. B. Curry, *Outlines of a Formalist Philosophy of Mathematics* (Studies in Logic and the Foundations of Mathematics), Amsterdam, North-Holland Publishing Co., 1951, pp. vii + 75, 7.50 fl.

S. Dasgupta, *A History of Indian Philosophy*, Vol. IV, *Indian Pluralism*, Cambridge at the University Press, 1949, pp. xiii + 483, £2 10s.

D. M. Datta, *The Chief Currents of Contemporary Philosophy*, Calcutta University, 1951, pp. xvi + 541, Rs. 10/8.

T. E. Davitt, *The Nature of Law*, London, B. Herder Book Co., 1951, pp. 274, £1 10s.

M. Deschoux, *Initiation à la Philosophie*, Paris, Presses Universitaires de France, 1951, pp. xii + 139, 320 frs.

K. Dürr, *The Propositional Logic of Boethius* (Studies in Logic and the Foundations of Mathematics), Amsterdam, North-Holland Publishing Co., 1951, pp. x + 79, 8 fl.

A. C. Ewing, *The Fundamental Questions of Philosophy*, London, Routledge & Kegan Paul, Ltd., 1951, pp. 260, 18s.

E. J. Furlong, *A Study in Memory*, A Philosophical Essay, London, Thomas Nelson & Sons Ltd., 1951, pp. 109, 12s. 6d.

S. Hampshire, *Spinoza*, Pelican Philosophy Series, 1951, pp. 235, 2s. 6d.

D. J. B. Hawkins, *The Meaning of Existentialism*, London, Blackfriars Publications, 1951, pp. 18, 1s. 6d.

L. Huovinen, *Das Bild Vom Menschen im Politischen Denken Niccolò Machiavellis*, Helsinki, Finnish Academy of Science and Letters, 1951, pp. 169, 400 m.

K. Jaspers, *La Situation Spirituelle de notre Epoque*, Louvain, E. Nauwelaerts, 1951, pp. 248, 69 Fr.b.

U. Klug, *Juristische Logik*, Berlin, Springer-Verlag, 1951, pp. 160, D.M. 18.60.

V. Kraft, *Die Grundlagen einer wissenschaftlichen Wertlehre*, Wien, Springer-Verlag, 1951, pp. 264.

B. B. v. F. g. Löringhoff, *Philosophical Problems of Mathematics* (translation), New York, Philosophical Library, 1951, pp. 88, $2.75.

W. D. Oliver, *Theory of Order*, Yellow Springs, Ohio, The Antioch Press, 1951, pp. x + 345, $5.00.

J. A. Passmore, *Ralph Cudworth*, An Interpretation, Cambridge at the University Press, 1951, pp. x + 120, 15s.

M. dal Pra, *Nicola di Autrecourt*, Milan, Fratelli Bocca Editori, 1951, pp. 195, L.700.

B. Rothenberg, *New Ways of Thinking* (in Hebrew) (Philosophical Inquiries, II), Tel Aviv, Israel, The Philosophical Society, 1949, pp. 106.

Sir C. Sherrington, *Man on his Nature*, second ed. (Gifford Lectures), London, Cambridge University Press, 1951, pp. 300, £1 10s.

L. Stefanini, *Esistenzialismo Ateo ed Esistenzialismo Teistico*, Padova, CEDAM, 1952, pp. 370, L. 1500.

St. Thomas Aquinas Philosophical Texts, Selected and translated by T.

Gilby, Oxford University Press (London: Geoffrey Cumberlege), pp. xxii + 405, 12s. 6d.

G. N. M. Tyrrell, *Homo Faber*, A Study of Man's Evolution, London, Methuen & Co. Ltd., 1951, pp. viii + 205, 15s.

Umanesimo e Scienza Politica, Edited by E. Castelli, Milan, Dott. Carlo Marzorati, 1951, pp. xiv + 570, L.3000.

W. M. Urban, *Humanity and Deity*, London, George Allen & Unwin Ltd., 1951, pp. 479, £1 5s.

W. J. Verdenius, *Mimesis, Plato's Doctrine of Artistic Imitation and its meaning to us*, Leiden, E. J. Brill, 1949, pp. 50, 2.30 guilders.

J. Wahl, *La pensée de l'existence*, Paris, Flammarion, 1951, pp. 290, 625 frs.

G. Williams, *Humanistic Ethics*, New York, Philosophical Library, 1951, pp. xii + 223, $3.75.

H. H. Williams, *Logic for Living*, New York, Philosophical Library, 1951, pp. xix + 281, $3.75.

G. H. von Wright, *An Essay in Modal Logic* (Studies in Logic and the Foundations of Mathematics), Amsterdam, North-Holland Publishing Co., 1951, pp. vi + 90, 9 fl.

H. Ahlenstiel, *Rotgrünblindheit als Erlebnis*, Göttingen, "Musterschmidt", Wissenschaftlicher Verlag, 1951, pp. 47.

L'Année Psychologique, by H. Piéron, A. Fessard and P. Fraisse, Vol. Jubilaire Hommage a Henri Piéron, Paris, Universitaires de France, 1951, pp. 711, 2400 frs.

A. Coates, *Prelude to History*, London, Methuen & Co. Ltd., 1951, pp. xviii + 289, £1 2s. 6d.

S. Daval and B. Guillemain, *Psychologie*, 2 vols, *Méthodes et Champs d'Applications* and *Les Fonctions Psychiques*, Paris, Presses Universitaires de France, 1951, pp. vii + 998, 900 frs. and 680 frs.

H. Delgado, *Introduccion à la Psicopatologia*, Buenos Aires, Ano del Libertador General San Martin, 1950.

J. M. Fraser, *Psychology: General, Industrial, Social*, London, Sir Isaac Pitman & Sons, Ltd., 1951, pp. x + 310.

K. H. Fremming, *The Expectation of Mental Infirmity in a sample of the Danish Population*, London, Cassell & Co. Ltd. (Occasional Papers on Eugenics, No. 7), 1951, pp. 53, 3s. 6d.

M. Gumpert, *The Anatomy of Happiness*, London, Hammond, Hammond & Co., Ltd., 1951, pp. 256, 15s.

L. R. Hubbard, *Dianetics*, The Modern Science of Mental Health, London, Derrick Ridgway, 1951, pp. xvii + 413, £1 10s.

G. Humphrey, *Thinking*, An Introduction to its experimental psychology, London, Methuen & Co. Ltd., 1951, pp. ix + 331, £1 1s.

W. Keller, *El Concepto Positivo de Existencia y la Psicologia*, Buenos Aires, Ano del Libertador General San Martin, 1950, pp. 27.

F. H. Lawson, *The Rational Strength of English Law*, London, Stevens & Sons Ltd., 1951, pp. viii + 147, 10s.

A. Rickel, *Eidetica*, Buenos Aires, Ano del Libertador General San Martin, 1950, pp. 54.

F. V. Smith, *The Explanation of Human Behaviour*, London, Constable & Co. Ltd., 1951, pp. ix + 276, 18s.

Martti Takala, *Asymmetries of the Visual Space*, Helsinki, The Finnish Academy of Science and Letters, 1951, pp. 175.

E. Utitz, *Caracterologia*, Buenos Aires, Ano del Libertador General San Martin, 1950, pp. 44.

C. Wolff, *The Hand in Psychological Diagnosis*, London, Methuen & Co. Ltd., 1951, pp. xv + 218, £1 12s. 6d.

Sri Aurobindo : The Prophet of Life Divine, by Haridas Chaudhuri, Calcutta Sri Aurobindo Pathamandir, 1951, pp. ix + 265, Rs. 3/-.

L. Babb, *Elizabethan Malady*, A study of Melancholia in English Literature from 1580 to 1640, Michigan State College Press (English agents : Basil Blackwell, Oxford), 1951, pp. ix + 206, £1 5s.

Rabbi B. Z. Bokser, *The Wisdom of the Talmud*, A Thousand Years of Jewish Thought, New York, Philosophical Library, 1951, pp. xx + 180, $3.75.

O. Cullmann, *Les Sacrements dans l'Evangile Johannique, La vie de Jésus et le culte de l'église primitive*, Paris, Presses Universitaires de France, 1951, pp. 92, 400 frs.

A. Dondeyne, *Foi Chrétienne et pensée Contemporaine*, Louvain, E. Nauwelaerts, 1951, pp. viii + 224, 100 Fr.b.

W. Empson, *The Structure of Complex Words*, London, Chatto & Windus, 1951, pp. 450, £1 1s.

A. D. Galloway, *The Cosmic Christ*, London, James Nisbet & Co. Ltd., 1951, pp. xiv + 274, 16s.

E. R. Hughes, *The Art of Letters*, Lu Chi's " Wen Fu ", A.D. 302, Translation and Comparative Study, Bollingen Series XXIX, Pantheon Books, pp. xviii + 261, $4.50.

J. H. Krimsky, *Jesus and the Hidden Bible*, New York, Philosophical Library, 1951, pp. vii + 81, $2.75.

R. B. Onians, *The Origins of European Thought*, About the Body, the Mind, the Soul, the World, Time and Fate. Cambridge at the University Press, 1951, £2 5s.

The Rockefeller Foundation Directory of Fellowship Awards for the years 1917–50, New York, 1951, pp. 286.

Srimat Swami, *Sri Sri Nityagopal*, Calcutta, Das Gupta & Co., 1948, Rs. 3/8.

R. M. T., *A Book of Preparation for the Coming Light*, London, Skeffington & Son Ltd., 1951, pp. 208, 9s. 6d.

P. Tillich, *The Protestant Era*, Trans. and ed. by J. L. Adams, London, Nisbet & Co. Ltd., 1951, pp. xlv + 305, £1 1s.

J. M. Warbeke, *The Power of Art*, New York, Philosophical Library, 1951, pp. ix + 493, $6.00.

XI.—NOTES

The Institute for the Unity of Science is offering a prize of $500 for the best essay on the theme "Mathematical Logic as a Tool of Analysis: Its Uses and Achievements in the Sciences and Philosophy". Two additional prizes of $200 each will be given for the next best two essays. It is an International Contest and is open to everyone. Essays must not exceed 25,000 words. They may be written in English, French or German and must be submitted before January 1, 1953. Further information can be obtained from the Institute for the Unity of Science, American Academy of Arts and Sciences, 28 Newbury Street, Boston 16, Massachusetts.

MIND ASSOCIATION

Those who wish to join the Association should communicate with the Hon. Treasurer, Mr. J. D. MABBOTT, St. John's College, Oxford, to whom the yearly subscription of sixteen shillings (payable in advance) should be sent. Cheques should be made payable to the Mind Association. Members may pay a Life Composition of £16 instead of the annual subscription. the annual subscription may be paid by Banker's Order; forms for this purpose can be obtained from the Hon. Treasurer.

In return for their subscriptions members receive MIND gratis and post free, and (if of 3 years' standing) are entitled to buy back numbers of both the Old and the New Series at half-price, if still in stock.

The Hon. Secretary of the Association is Professor KARL BRITTON, Dep. of Philosophy, King's College, Newcastle-on-Tyne.

Members resident in U.S.A. may pay the subscription ($2.80) to the Hon. Assistant-Treasurer, Prof. B. Blanshard, Dept. of Phil., Yale University, New Haven, Conn.

JOINT SESSION OF THE ARISTOTELIAN SOCIETY AND THE MIND ASSOCIATION AT BIRMINGHAM

JULY 11th to 13th, 1952

Programme

FRIDAY, JULY 11th

At 8 p.m. Chairman: Professor A. J. Ayer. Presidential Address by Professor A. E. Duncan Jones. "Deadlock in Ethics."

SATURDAY, JULY 12th

At 10 a.m. Chairman: Professor C. D. Broad. Symposium: "Seeming". Professor Karl Britton; Professor H. H. Price; Mr. A. M. Quinton.

At 8 p.m. Chairman: Professor L. J. Russell. Symposium: "Men and Machines". Mr. J. O. Wisdom; Mr. D. M. Mackay; Mr. R. J. Spilsbury.

SUNDAY, JULY 13th

10 a.m. Chairman: Mr. A. R. MacIver. Symposium: "Reducibility". Mr. J. F. Thompson; Mr. G. J. Warnock; Mr. R. B. Braithwaite.

At 8 p.m. Chairman: Mr. J. L. Austin. Symposium: "Motives and Causes". Mr. R. Peters; Mr. D. J. McCracken; Mr. J. O. Urmson.